

МАТЕРИАЛЫ К ИСТОРИИ АКАДЕМИИ НАУК СССР

(1917—1947)




1917
1947


АКАДЕМИЯ
НАУК
СССР

АКАДЕМИЯ НАУК
Союза
Советских
Социалистических
Республик



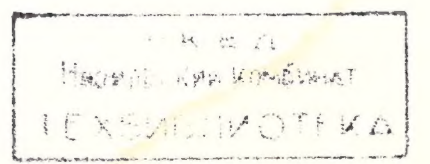
ЮБИЛЕЙНЫЙ СБОРНИК

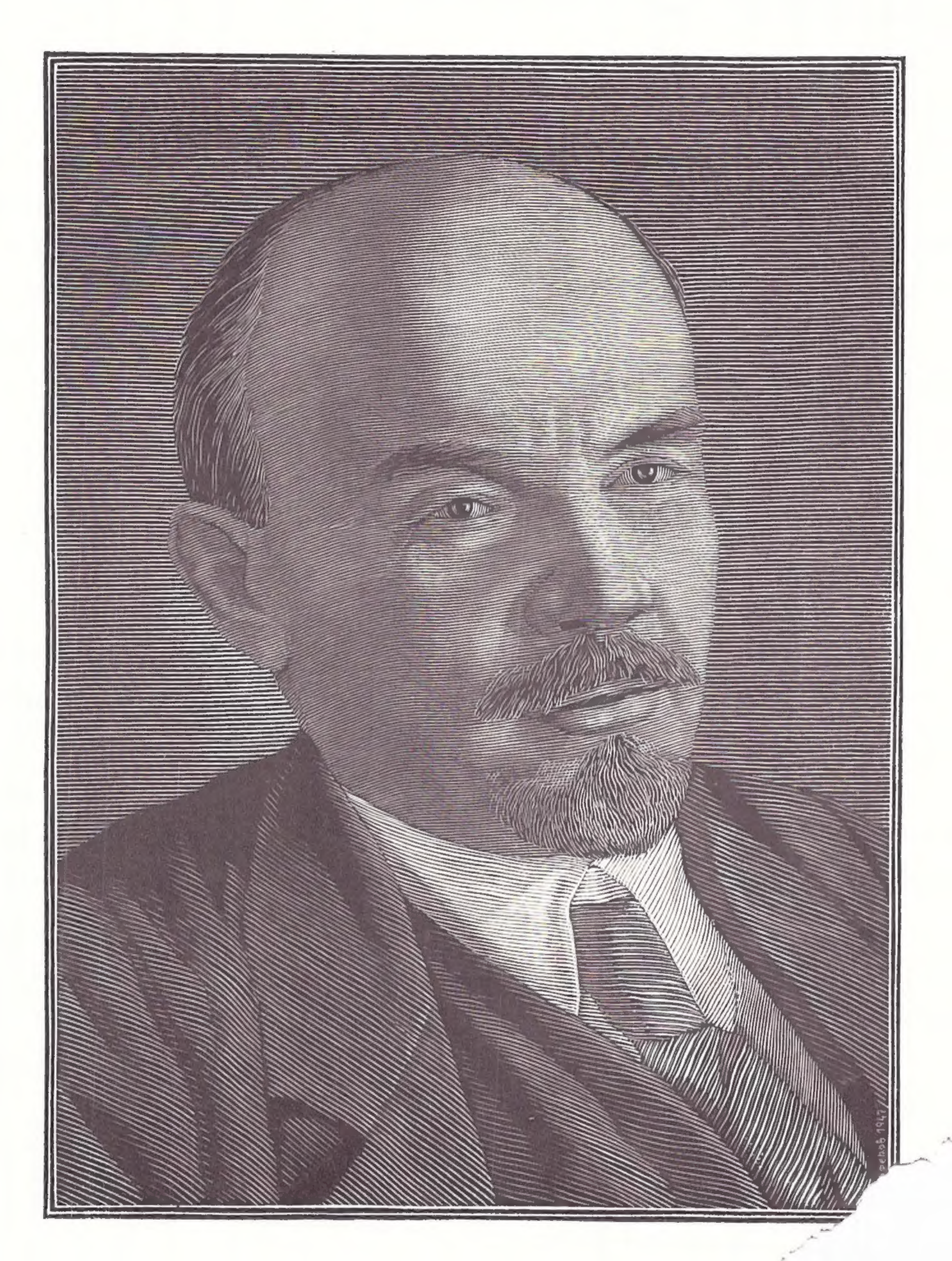
ПОСВЯЩЕННЫЙ ТРИДЦАТИЛЕТИЮ ВЕЛИКОЙ ОКТЯБРЬСКОЙ СОЦИАЛИСТИЧЕСКОЙ РЕВОЛЮЦИИ

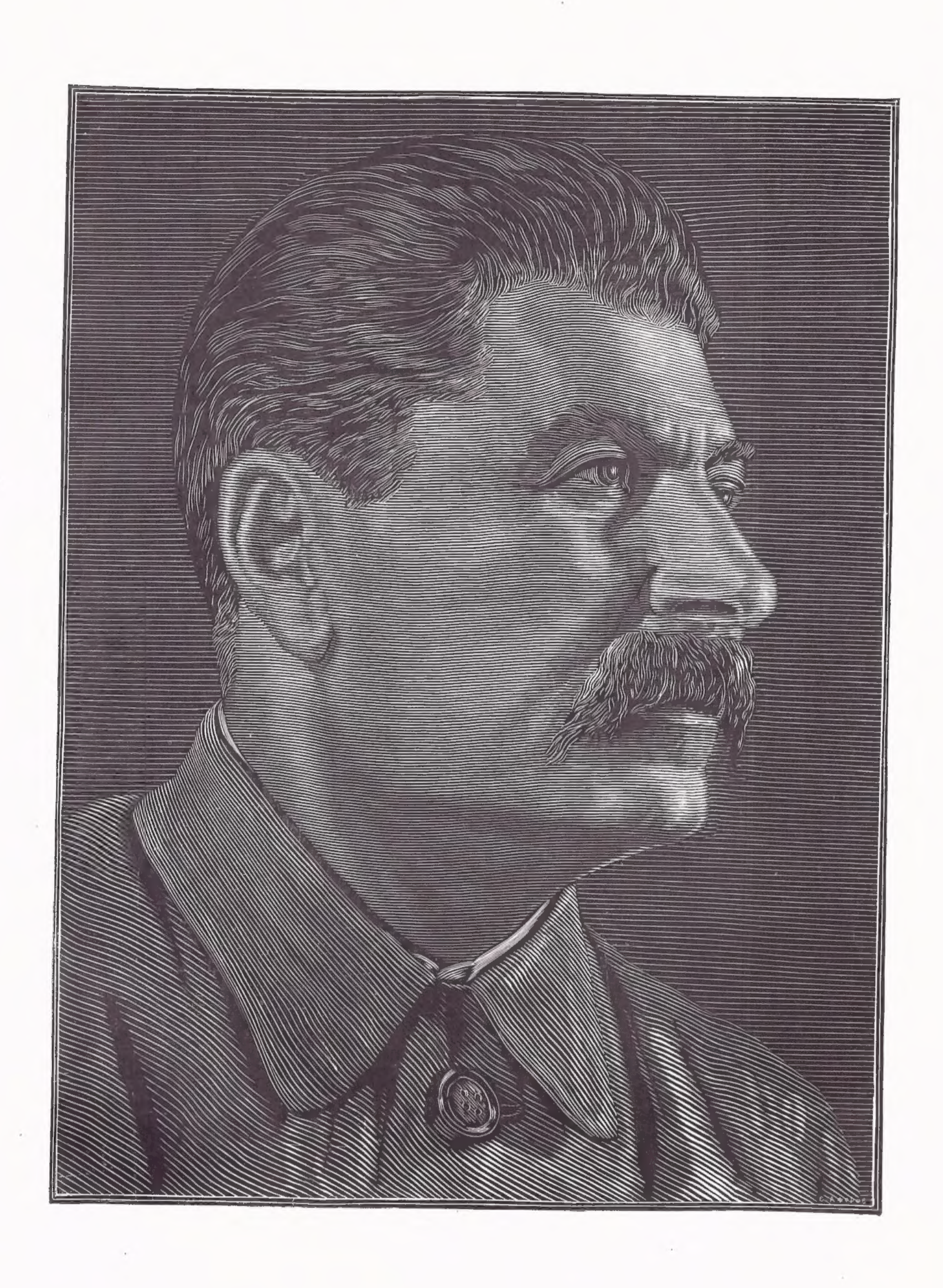


В ДВУХ ЧАСТЯХ

ИЗДАТЕЛЬСТВО АКАДЕМИИ НАУК СССР
МОСКВА · ЛЕНИНГРАД
1 9 4 7







ПРЕДИСЛОВИЕ

Академия Наук СССР посвящает эту книгу тридцатой годовщине важнейшего события в истории нашей Родины и всего человечества — Великой Октябрьской социалистической революции.

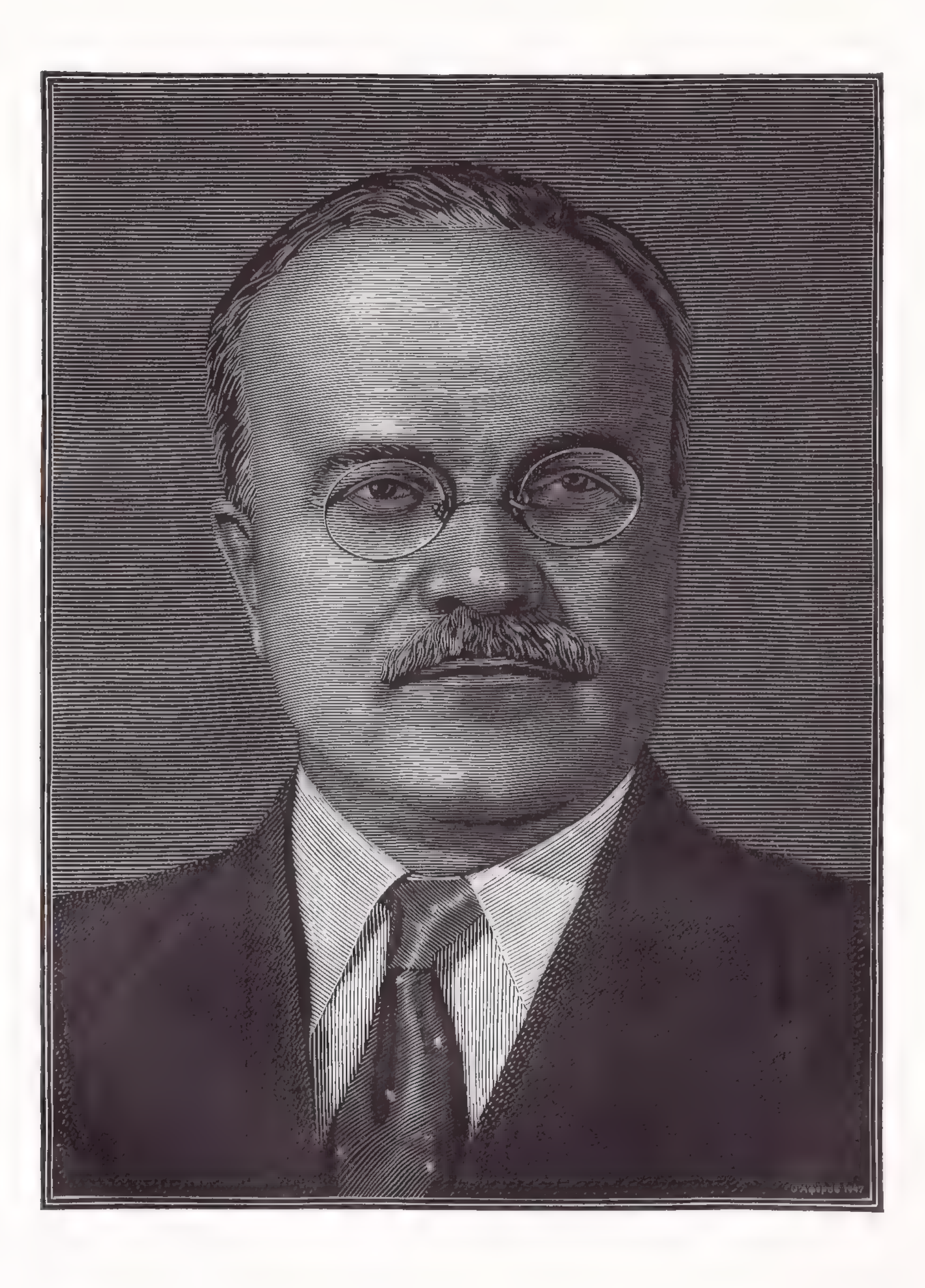
Огромны и неисчислимы следствия Октябрьской революции. Нет такой стороны жизни, на которую они не повлияли бы коренным и решающим образом. Самое глубокое изменение произошло в людях, в обществе, образующем наше государство. Дореволюционная Российская империя была государством с резким классовым расслоением, с узаконенными вековыми социальными несправедливостями, основанными на эксплоатации труда рабочих, крестьян и интеллигенции, на бесправии национальных меньшинств. За тридцать лет напряженного труда всего народа под руководством Коммунистической партии и Советского правительства бывшая Российская империя превращена в государство нового типа, в сплоченный, многонациональный Союз Советских Социалистических Республик, в дружную социалистическую семью народов.

Из жизни советского общества устранены классовые преграды, исчезли эксплоатация, обогащение одного человека за счет других, нет больше национальных ограничений, нет неравенства в правах мужчин и женщин. В соответствии с этим неузнаваемо преобразилась духовная жизнь советского человека, изменилась его психология. Доминирующим для каждого стало чувство общественного долга.

Небывалая в истории человечества по своим следствиям и глубине социальная революция сопровождалась крутым переломом во всех проявлениях культурной жизни страны. Резко ускорился рост нашей науки, во многом изменилось ее содержание, наука чрезвычайно дифференцировалась благодаря новой роли, которую она приобрела в социалистическом государстве.

Петр I создал необходимые первоначальные условия для роста науки в нашей стране. С тех пор русская наука начала свое поступательное движение. Научное творчество таких русских ученых, как Ломоносов, Лобачевский, Менделеев, Павлов и многие другие, показало всему миру научный потенциал нашей Родины.

Однако и после Петра роль науки в России оставалась ограниченной. Наука требовалась государству главным образом для подготовки учителей, профессоров, инженеров, для освоения мировой науки. Научное творчество, широкое продвижение собственных исследований, своя наука, ломающая традиции и устанавливающая новые вехи для понимания природы и для развития техники, не получали от государства поддержки и новых заданий. Русские великие ученые в большинстве случаев были одиночками, не имели школ и последователей, не располагали необходимым научным оборудованием и средствами. Поэтому многие важные научные результаты,



достигнутые русскими учеными, открытия и изобретения, сделанные в России до революции, далее не развивались и нередко использовались сначала за границей.

Великая Октябрьская социалистическая революция положила этому конец. С первых же месяцев существования советской власти научно-исследовательская работа получила со стороны Коммунистической партии и Советского правительства поддержку, невиданную ранее. В Москве, Ленинграде и ряде других городов стали возникать большие исследовательские институты по различным отраслям наук. Академия Наук из небольшого и замкнутого научного учреждения за тридцать лет преобразовалась в огромную, едва ли не самую большую в мире ассоциацию исследовательских институтов, лабораторий и станций. В настоящее время число сотрудников Академии Наук более чем в 40 раз превосходит состав дореволюционной Академии. Филиалы и базы Академии распределены по всей стране, от Кишинева и Петрозаводска до Якутска и Южного Сахалина.

Наряду с Академией Наук СССР учреждены и ведут очень большую работу другие академии: 10 Академий Наук союзных республик — Украинской, Белорусской, Грузинской, Азербайджанской, Армянской, Узбекской, Казахской, Литовской, Латвийской, Эстонской и 5 академий отраслевого характера — Академия сельскохозяйственных наук имени В. И. Ленина, Академия архитектуры, Академия медицинских наук, Академия педагогических наук, Академия артиллерийских наук. Эти 15 академий имеют много специальных институтов и в целом составляют громадную и разнообразную научно-исследовательскую систему. Исследовательская работа, непосредственно связанная с запросами промышленности и техники, в широких размерах ведется также в больших специальных институтах. Таковы Центральный Аэрогидродинамический институт,

Государственный Оптический институт, Всесоюзный Электротехнический институт, Институт имени Карпова, Центральный Геологоразведочный институт и многие другие.

Высшие школы СССР, по хорошей традиции, наряду с основной работой по подготовке кадров ведут очень большую исследовательскую работу во всех областях знаний.

В общем к делу научного исследования в СССР привлечено около сотни тысяч людей; до революции таких работников было приблизительно в 50 раз меньше.

Чрезвычайно резкий подъем исследовательской работы, прокладывающей новые пути, — одна из самых характерных черт, отличающих советскую науку от дореволюционной. Самое содержание науки во многом решительно изменилось; это особенно характерно для гуманитарных наук. Методологической базой всех наук является у нас диалектический и исторический материализм, учение Маркса и Энгельса, развитое Лениным и Сталиным. Совершенно изменились и методы и выводы исторического исследования. Правовая и экономическая наука советской страны совсем не похожа на учение о праве и экономику дореволюционной России. Характерные изменения испытали также естественнонаучные дисциплины и техника. Каждая наука служит обществу, народу. Отсюда — своеобразный «практицизм» советской науки, не порывающей, одновременно, даже в самых детальных, конкретных своих выводах связи с теорией, с широкими обобщениями.

В советской стране наука стала народной не только потому, что она постоянно служит интересам народа, но и потому, что она максимально доступна для народа. Кадры современных советских молодых ученых составились по преимуществу из рабочих и крестьян. Ученые — уже не изолированная каста, они слились с народом.

Большие и ответственные задачи решались советской наукой и техникой в годы сталинских пятилеток. Великая Отечественная война была серьезным испытанием для нашей науки и техники. Это испытание выдержано с честью. Новая наша наука во многом помогла успеху дела построения бесклассового социалистического общества и навеки памятной победе над фашизмом.

☆

В прежние годы Академия Наук СССР много раз издавала специальные книги, посвященные истории и итогам советской науки в различных областях. Настоящий сборник статей академиков и других ученых, работающих в Академии, отличается от прежних юбилейных изданий. Его задача — дать некоторые примеры достижений нашей науки за тридцать советских лет в статьях, написанных авторами научных исследований или их учениками. Сборник, конечно, не может показать все важнейшие результаты советской науки даже в рамках одной Академии.

Велика, необычайно разнообразна и с большим трудом обозрима современная наука, поднявшаяся на плодородной почве могучей социалистической страны. Нет сомнения, что этот процесс роста и дифференциации не остановится, а, наоборот, умножится в ближайшие годы.

9 февраля 1946 г. И. В. Сталин поставил перед советскими учеными прекрасную цель — догнать и превзойти в науке достижения зарубежных стран. Этого требуют величественные задачи, стоящие перед СССР.

ЧАСТЬ ПЕРВАЯ



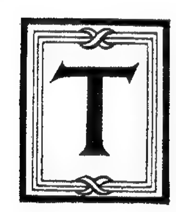
Академик

А. Я. Вышинский

УЧЕНИЕ ЛЕНИНА—СТАЛИНА О ПРОЛЕТАРСКОЙ РЕВОЛЮЦИИ И ГОСУДАРСТВЕ

I

ВЕЛИКАЯ ОКТЯБРЬСКАЯ СОЦИАЛИСТИЧЕСКАЯ РЕВОЛЮЦИЯ И ПОСТРОЕНИЕ СОЦИАЛИЗМА В ОДНОЙ СТРАНЕ



ридцать лет назад, 25 октября (7 ноября) 1917 г., рабочие и крестьяне России под руководством большевистской партии Ленина — Сталина свергли власть помещиков и капиталистов и взяли власть в свои руки. Победила Великая Октябрьская социалистическая революция. Открылась новая страница в истории нашей родины. Победа Октябрьской революции означала «коренной перелом в истории человечества, коренной перелом в исторических судьбах мирового капитализма, коренной перелом в освободительном движении мирового пролетариата, коренной перелом в способах борьбы и формах организации, в быту и традициях, в культуре и идеологии эксплоатируемых масс всего мира».[1]

В своей теории пролетарской революции Ленин исходил из открытого им закона неравномерного развития капитализма. Ленин сформулировал этот закон, опираясь в полной мере на установленные Марксом и

[1] Сталин. Вопросы ленинизма, изд. 11-е, стр. 175.

Энгельсом законы капиталистического развития и, в первую очередь, на марксистский тезис о том, что «свободная конкуренция порождает концентрацию производства, а эта концентрация на известной ступени своего развития ведет к монополии».[1]

Ленин, в свою очередь, показал, что «порождение монополии концентрацией производства вообще является общим и основным законом современной стадии развития капитализма».[2]

На анализе многочисленных данных буржуазных экономистов Ленин показал, до какой степени выросла банковская монополия, как «количество перешло в качество», в чем именно это выразилось, как осуществляется переход развитого капитализма в империализм.

Ленин показал, — и это представляет громадный интерес и имеет огромное значение и в настоящее время, т. е. спустя 33 года после того, как это было впервые написано Лениным в его бессмертной работе «Империализм, как высшая стадия капитализма», — что «империализм есть эпоха финансового капитала и монополий, которые всюду несут стремления к господству, а не к свободе» и что результатом этих тенденций является «реакция по всей линии при всяких политических порядках, крайнее обострение противоречий и в этой области».[3]

Ленин цитировал Гильфердинга, которого нельзя заподозрить в радикализме, утверждавшего, что в эпоху империализма и под влиянием ввозимого во вновь открытые страны капитала «в корень революционизируются старые социальные отношения, разрушается тысячелетняя аграрная обособленность «вне-исторических наций», они вовлекаются в капиталистический водоворот. Сам капитализм мало-по-малу дает покоренным средства и способы для освобождения. И они выдвигают ту цель, которая некогда представлялась европейским нациям наивысшею: создание единого национального государства, как орудия экономической и культурной свободы. Это движение к независимости угрожает европейскому капиталу в его наиболее ценных областях эксплуатации, сулящих наиболее блестящие перспективы, и европейский капитал может удерживать господство, лишь постоянно увеличивая свои военные силы».[4]

---

[1] Ленин. Соч., т. XIX, стр. 84.
[2] Там же, стр. 85.
[3] Там же, стр. 169.
[4] Там же.

Ленин добавил к сказанному Гильфердингом, что не только во вновь открытых, но и в старых странах империализм ведет к аннексии, к усилению национального гнета и, следовательно, также к обострению сопротивления.

К какому выводу должен был привести этот анализ? К такому, что в условиях империализма развитие вперед «осуществимо лишь к *социалистическому* обществу, к *социалистической революции*».[1] К такому именно выводу пришел Ленин. Этот вывод имеет чрезвычайно важное значение. Он говорит, во-первых, о том, что империализм неизбежно ведет к усилению реакции, выражающейся в бесцеремонном усилении нажима на более слабые в экономическом и военном отношении страны. Жажда наживы и господства обнажается до предела. Спадает вуаль демократии, скрывающая голую империалистическую алчность, и оскал хищника предстает перед всеми в своем откровенном виде.

Он говорит далее о том, что обостряется стремление к аннексиям, т. е., как подчеркивал Ленин, к нарушениям национальной независимости. Он говорит, наконец, о неизбежно возрастающем сопротивлении этим тенденциям, о революционизировании старых социальных отношений, об усугубляющейся среди угнетенных наций тенденции встать на защиту своей независимости.

Современная стадия развития империализма полностью подтверждает этот гениальный анализ и гениальный вывод Ленина. Доказательством безупречной правильности такого вывода может служить внешнеполитическое положение Соединенных Штатов Америки, представляющих собой в настоящее время наиболее характерный образец государства, где господствует финансовый капитал и где курс внешней политики полностью отражает указанные выше черты.

США оказывают «помощь» Греции и Турции. США готовы оказать «помощь» Франции, при том, впрочем, условии, если французские «демократы» и «социалисты» согласятся изгнать из правительства коммунистов и не пускать их обратно. США готовы «помочь» Великобритании, если лейбористское правительство откажется от Рура и от своего плана национализации. Они готовы оказать «помощь» Италии, если де Гаспери и Ватикан тоже сумеют изгнать из правительства коммунистов и

[1] Ленин. Соч., т. XIX, стр. 187.

создать такое правительство, которое было бы готово поклоняться святому... доллару.

США готовятся «помогать» Иемену, Ирану, Ираку и вообще любой стране, готовой, в свою очередь, опереться на американские финансы и американскую атомную бомбу.

На наших глазах совершается, таким образом, жертвоприношение на алтарь американского доллара, единственного властителя человеческих судеб в мире империализма. Это жертвоприношение совершается в соответствии с новейшей американской процедурой, действующей по всем правилам патентованной американской демократии.

40 лет назад, говоря о мировой политике империалистических стран, Ленин писал: «Сознательные вожди империализма говорят себе: мы не можем, конечно, осуществить свои цели без удушения мелких народов, но ведь есть два способа удушения. Бывают случаи, когда надежнее — и выгоднее — получить искренних, добросовестных «защитников отечества» в империалистской войне путем создания *политически* независимых государств, о *финансовой* зависимости которых «мы» уже позаботимся!»[1]

Кажется, что эти строки написаны только сейчас, — настолько они свежи и настолько метко они бьют в самое больное место современного империализма!

Анализируя развитие противоречий в системе империализма, Ленин указывал на то, что «капитализм перерос во всемирную систему колониального угнетения и финансового удушения горстью «передовых» стран гигантского большинства населения земли», что «дележ этой «добычи» происходит между 2—3 всемирно могущественными, вооруженными с ног до головы хищниками (Америка, Англия, Япония), которые втягивают в *свою* войну из-за дележа *своей* добычи всю землю».[2]

Ленин указывал на то, что «рост противоречий внутри мировой системы финансового угнетения и неизбежность военных столкновений ведут к тому, что мировой фронт империализма становится легко уязвимым со стороны революции, а прорыв этого фронта со стороны отдельных стран—вероятным».[3]

Отсюда Ленин приходил к двум выводам, имеющим всемирно-истори-

[1] Ленин. Соч., т. XIX, стр. 383.
[2] Там же, стр. 74.
[3] Сталин. Вопросы ленинизма, стр. 85—86.

ческое значение и определившим собой, вместе с тем, тактику русских коммунистов в преддверии Октября 1917 г.:

а) Прорыв фронта империализма «вероятнее всего может произойти в тех пунктах и в тех странах, где цепь империалистического фронта слабее, т. е. где империализм менее всего подкован, а революции легче всего развернуться».[1]

б) «Ввиду этого победа социализма в одной стране, если даже эта страна является менее развитой капиталистически, при сохранении капитализма в других странах, если даже эти страны являются более развитыми капиталистически, — вполне возможна и вероятна».[2]

Таковы, говорит товарищ Сталин, в двух словах основы ленинской теории революции. Эта теория революции имеет мировое значение, так как она явилась результатом подлинно научного анализа империализма, путей его развития и его гибели. Ленинская теория революции — творческая и созидательная, зовущая к действию, организующая и направляющая. В противоположность теории оппортунистов, выжидающих, когда созреет пролетарская революция в наиболее развитой капиталистической стране, ленинская теория говорит о пролетарской революции там, где ситуация складывается более революционно, хотя бы эта страна в капиталистическом отношении была менее развитой. Это — новая постановка вопроса, явившаяся результатом дальнейшего развития Лениным основных принципов марксистской теории пролетарской революции. Ленин и Сталин, двигая марксистскую науку вперед, показали, что в условиях империализма уже оказывается недостаточной точка зрения, исходящая из наличия или отсутствия объективных условий пролетарских революций в отдельных странах, или точнее — в той или иной стране. «Теперь, — учит Сталин, — нужно говорить о наличии объективных условий революции во всей системе мирового империалистического хозяйства, как единого целого...».[3] «Теперь, — учит Сталин, — надо рассматривать пролетарскую революцию, прежде всего, как результат развития противоречий в мировой системе империализма, как результат разрыва цепи мирового империалистического фронта в той или иной стране».[4]

[1] Сталин. Вопросы ленинизма, стр. 86.
[2] Там же.
[3] Там же, стр. 18.
[4] Там же, стр. 19.

Раньше, т. е. в эпоху доимпериалистической стадии капитализма, считали, что пролетарская революция начнется там, где больше развита промышленность, где пролетариат составляет большинство. «Нет, — возражает ленинская теория революции, — *не обязательно там, где промышленность больше развита*, и пр. Фронт капитала прорвется там, где цепь империализма слабее, ибо пролетарская революция есть результат разрыва цепи мирового империалистического фронта в наиболее слабом ее месте...».[1]

Так случилось в России в 1917 г., где «цепь империалистического мирового фронта оказалась слабее», чем в других странах. «Там, — говорит Сталин, — она и прорвалась, дав выход пролетарской революции».

Причина этого лежала в том, что «в России развертывалась величайшая народная революция, во главе которой шел революционный пролетариат, имевший такого серьезного союзника, как многомиллионное крестьянство, угнетаемое и эксплоатируемое помещиком».[2] Причина была в том, что «против революции стоял там такой отвратительный представитель империализма, как царизм, лишенный всякого морального веса и заслуживший общую ненависть населения. В России цепь оказалась слабее, хотя Россия была менее развита в капиталистическом отношении, чем, скажем, Франция или Германия, Англия или Америка».[3]

Таким образом, на долю тогдашней России выпало прорвать цепь мирового империализма и начать эру пролетарских революций.

Ленин и Сталин, возглавившие пролетарскую революцию в 1917 г. в России, опирались в своей уверенности на успех восстания, на анализ сложившейся тогда обстановки в нашей стране, на строгий и точный учет соотношения классовых сил. Они опирались на ленинскую теорию пролетарской революции, в основе которой лежит ленинский закон неравномерного развития капитализма и возможности победы социализма в нескольких или даже в одной, отдельно взятой стране. Это, как небо от земли, отлично от упаднической и демобилизующей троцкистской теории, исходившей из того, что пролетарская революция может победить лишь в международном масштабе и что победа социализма в одной стране невозможна.

---

[1] Сталин. Вопросы ленинизма, стр. 19.

[2] Там же.

[3] Там же.

Победа Октябрьской революции показала ошибочность универсальной теории одновременной победы революции в основных странах Европы, теории невозможности победы социализма в одной стране.[1] Известно, что именно эту теорию троцкисты выдавали за последнее слово науки, утверждая, что победа социализма в одной стране невозможна. Больше того, Троцкий прямо утверждал, что если бы «наша инициатива» в борьбе за социализм не дала толчка борьбе в других странах, то «безнадежно думать..., что, например, революционная Россия могла бы устоять перед лицом консервативной Европы, или социалистическая Германия могла бы остаться изолированной в капиталистическом мире».[2]

Троцкий противопоставлял ленинской теории победы социализма в одной стране теорию одновременной победы социализма в основных странах Европы. Троцкий и его подголоски своей «теорией» просто прикрывали измену делу социализма. Этой «теорией» они пытались обезоружить революционное социалистическое движение в нашей стране, парализовать его усилия, заковать его наглухо, остановить процесс социалистического строительства в СССР, демобилизовать и дезорганизовать ряды строителей социализма. Сначала мы принимали эту «т е о р и ю», как теорию, как идеологическое построение, хотя и принципиально расходящееся с нашей программой, но продиктованное чистыми, научными побуждениями. Впоследствии оказалось иное — было установлено, что троцкисты, зиновьевцы, бухаринцы, каменевцы и их единомышленники выполняли социальный заказ буржуазии и ее разведывательных органов, как агентура, продавшаяся врагам рабочего дела, врагам социализма.

С исключительным бесстыдством троцкисты разыгрывали порученную им роль «ученых теоретиков» социализма, подводя под свою писанину «теоретический» фундамент. Троцкий, например, не стыдился в обоснование своего утверждения, что революционная (т. е. социалистическая) Россия не могла бы устоять против консервативной Европы, ссылаться на «опыт истории» и «теоретические соображения». Однако Троцкому хорошо было известно, что «опыт истории» нельзя брать абстрактно, что история совершается в конкретных условиях и

---

[1] С т а л и н. Вопросы ленинизма, стр. 102.

[2] Цит. по книге: С т а л и н. Вопросы ленинизма, стр. 88.

что именно эти условия и определяют содержание и значение «опыта истории».

«Опыт истории», если взять опыт Советского государства, накопленный к тому времени (к 1922 г.), когда Троцкий особенно решительно выступал против возможности построения в СССР социализма, полностью свидетельствовал не против, а в пользу возможности построения социализма в СССР. Но это не смущало троцкистов, твердивших свое, каркавших о провале социалистического строительства в СССР.

Итоги первого советского пятилетия ознаменовались яркими доказательствами устойчивости советской власти, отбившей интервенцию 14 государств, подавившей восстания генеральско-помещичьей контрреволюции, во главе с эсерами и меньшевиками, закрепившей достижения Октября, блестяще осуществившей переход от военного коммунизма к новой экономической политике, явившейся новой исходной точкой дальнейшего социалистического наступления.

«Опыт истории», на который пытались ссылаться троцкисты, говорит против этих незадачливых «теоретиков», извративших марксизм, опошливших революционное учение Маркса — Энгельса о государстве и революции, о борьбе классов, о перспективах и условиях организации социалистического общества.

Когда СССР перешел к нэпу, эти «теоретики» из так называемой «новой оппозиции» объявили, что с социализмом у нас покончено, что нэп — это возврат к капитализму. Зиновьев, как всем памятно, сочинил даже по этому поводу свою пресловутую «философию эпохи», весь смысл которой заключался в проповеди полного неверия в победу социалистического строительства в нашей стране, неверия в способность нашего пролетариата повести за собой крестьянство по пути к социализму, неверия в социалистический путь развития крестьянства. Мы все помним, как Троцкий предсказывал, что пролетариат после победоносной революции неизбежно придет в столкновение с крестьянством, а ведь это писалось в 1922 г.! Он утверждал, что мы к созданию социалистического общества не пришли и даже не подошли и что «подлинный подъем социалистического хозяйства в России станет возможным *только после победы* пролетариата в важнейших странах Европы».[1]

---

[1] Цит. по книге: Сталин. Вопросы ленинизма, стр. 89.

Товарищ Сталин разоблачил эту, с позволения сказать, «теорию», как разновидность меньшевизма, как «перманентную» безнадежность и беспросветность, как неверие в силы и способности пролетариата России.

Попытки троцкистов противопоставить теорию «перманентной революции» подлинной революционной ленинской теории пролетарской революции позорно провалились. Беспощадная, опирающаяся на научно безупречную марксистско-ленинскую методологию, критика Сталиным троцкистских извращений марксизма-ленинизма вскрыла предательский характер троцкистско-зиновьевской «философии». В деле разоблачения троцкистских, зиновьевских и бухаринских лжеучений выдающуюся роль сыграли такие работы Сталина, как «Октябрьская революция и тактика русских коммунистов» (1924), «К вопросам ленинизма» (1926), «О правом уклоне в ВКП(б)» (1929) и др.

Сейчас, после тридцати лет победоносного социалистического строительства в СССР, когда остались позади основные трудности организации социалистического хозяйства и социалистических общественных отношений, когда решена задача построения в нашей стране социализма, снят самой жизнью вопрос о том, насколько правильна ленинская теория пролетарской революции.

История тридцатилетней борьбы за социализм в СССР, увенчавшаяся блестящей победой социализма на основе ленинско-сталинской теории научного социализма, говорит сама за себя и не нуждается в признании, как не нуждается в признании молодое, но могучее и великое Советское государство.

Великая Октябрьская революция явилась животворным процессом развития в тогдашней России, а затем в Союзе Советских Социалистических Республик новых общественных отношений. Гениальное предвидение Ленина и Сталина вывело тогдашнюю Россию на новый путь развития, поставив перед нашим народом гигантские задачи — довести до конца буржуазную революцию и осуществить одновременно революцию социалистическую, построив социалистическое общество в условиях капиталистического окружения.

Эти задачи встали во весь рост перед партией Ленина — Сталина, возглавившей в 1917 г. народные массы. Партия и народы России смело взялись за решение этих задач.

Что означало довести буржуазную революцию до конца? Это блестяще разъяснил Сталин в ответе Ян—скому в статье «О трех основных лозунгах партии по крестьянскому вопросу» (1927).

Вот что писал тогда Сталин: «Доведение до конца буржуазной революции не есть единичный акт. На деле оно растянулось на целый период, захватывая не только кусочки 1918 года, как вы (т. е. Ян—ский.—*А. В.*) утверждаете в своем письме, но и кусочки 1919 года (Поволжье — Урал) и 1919—1920 гг. (Украина). Я имею в виду наступление Колчака и Деникина, когда перед крестьянством в целом встала опасность восстановления помещичьей власти и когда оно, именно как *целое*, вынуждено было сплотиться вокруг Советской власти для того, чтобы обеспечить доведение до конца буржуазной революции и сохранить за собой плоды этой революции».[1]

Эта заинтересованность крестьянства в целом в успешном разрешении задач буржуазной революции и в завершении этой революции, что полностью обеспечивала пролетарская диктатура, создавала известную сложность, то, что Сталин охарактеризовал как «причудливое» переплетение непосредственно социалистических задач диктатуры с задачей доведения до конца буржуазной революции.

Этой особенности Октябрьской революции не поняли троцкисты и каменевцы, не поняли и ученые-историки из школки М. Н. Покровского, изображавшие дело так, будто «основным лозунгом» Октября был не переход государственной власти в руки пролетариата, не установление диктатуры пролетариата, а именно доведение до конца буржуазной революции.

То, что диктатура пролетариата «вымела дочиста грязь средневековья»,[2] ученые мужи типа проф. М. Н. Покровского выдавали за «основной лозунг», основную задачу Октября!

Дело обстояло, как мы видим, не так. Дело заключалось в своеобразном переплетении в ходе Октябрьской революции «непосредственно социалистических задач диктатуры с задачей доведения до конца буржуазной революции»[3] (подчеркнуто мною. — *А. В.*).

Но задача «доведения до конца буржуазной революции» была

[1] Сталин. Вопросы ленинизма, стр. 160.

[2] Там же.

[3] Там же.

«побочным» продуктом Октябрьской революции, разрешившей эту задачу «мимоходом».[1]

Главное же было в свержении власти буржуазии, в переходе власти в руки пролетариата, в организации социалистического государственного и общественного строя. Октябрьская революция начисто вымела дворянско-помещичий мусор и уже этим подвигом своим вошла в историю человечества как Великая революция. Но Октябрьская революция еще более велика тем, что она явилась социалистической революцией.

«Отныне наступает новая полоса в истории России, и данная третья русская революция должна в своем конечном итоге привести к победе социализма».[2]

Эти слова Ленина явились подлинно историческими, ибо в них отразилась вся сила уверенности вождя революции в победоносном ее исходе, уверенности в том, что, завершив этап буржуазно-демократического революционного развития, пролетарская революция поднимается на новую, высшую ступень своего движения вперед, к социализму. Эта уверенность явилась могучей движущей силой и в последующую эпоху завершения социалистического строительства в СССР под гениальным руководством товарища Сталина.

Сила этой уверенности определялась научным предвидением Ленина и Сталина, в совершенстве овладевших могучим оружием диалектического анализа и синтеза. Это оружие открыло перед умственным взором гениальных вождей пролетариата значение прошлого и перспективы будущего. Это предвидение Ленина и Сталина и их ближайших соратников из Центрального Комитета большевистской партии обеспечило в Октябре 1917 г. правильный выбор момента для нанесения главного удара.

Напомним, что 24 октября (6 ноября) в письме членам ЦК Ленин требовал немедленного захвата власти. «Я пишу эти строки, — читаем мы в этом историческом документе, — вечером 24-го, положение донельзя критическое. Яснее ясного, что теперь, уже поистине, промедление в восстании смерти подобно».[3]

И дальше: «Нельзя ждать!! Можно потерять все!!.. История не простит промедления революционерам, которые могли победить сегодня

[1] Сталин, Вопросы ленинизма, стр. 172.
[2] Ленин. Соч., т. XXII, стр. 4.
[3] Там же, т. XXI, стр. 362.

(и наверняка победят сегодня), рискуя терять много завтра, рискуя потерять все».[1]

Ленин требовал немедленного выступления, говоря, что история всех революций доказала, что «было бы гибелью или формальностью ждать колеблющегося голосования 25 октября», что «народ вправе и обязан решать подобные вопросы не голосованиями, а силой», что «народ вправе и обязан в критические моменты революции направлять своих представителей, даже своих лучших представителей, а не ждать их».[2]

Только гений революции мог так говорить, ощущая всем своим существом неизбежность приближающейся победы, чувствуя и своим умом и своим сердцем биение сердца народа, ждущего призыва своего вождя, чтобы ринуться в последний и решительный бой!..

История еще раз доказала, что Ленин был прав, что был прав Сталин, была права партия. Ход событий показал, что успех Октябрьской революции зависел не только от тех общеполитических условий, которые были непосредственно вызваны мировой войной. Успех зависел в решающей степени от того, что в бой против русских капиталистов и помещиков вступила закаленная в классовых битвах армия рабочих и крестьян под руководством большевистской партии Ленина — Сталина, что буржуазия, возглавленная истериком Керенским, эсерами и меньшевиками вкупе с контрреволюционными генералами и адвокатами, оказалась неспособной выдержать удар питерских и московских рабочих, противостоять революционной буре, прокатившейся могучими раскатами по всей стране. Успех зависел от того, что так правильно был выбран и с математической точностью нанесен этот удар, решивший судьбу революции.

Ленин и Сталин правильно оценили сложившуюся в то время ситуацию, правильно поняли ход событий, правильно определили линию развития этих событий, мчавшихся навстречу революционной разрядке, к своему логическому завершению. Прочтите ленинские статьи кануна Октября: «Удержат ли большевики государственную власть?», «Кризис назрел», «Задачи революции», «Письмо к товарищам большевикам», «Письмо к товарищам» и др. Прочтите сталинские статьи более позднего времени: «Октябрьская революция и тактика русских коммунистов», «К вопросам

---

[1] Ленин. Соч., т. XXI, стр. 362.

[2] Там же, стр. 363.

ленинизма», «О трех основных лозунгах партии», где Сталин дает сокрушительный отпор троцкистским попыткам извратить историю Октября, с исключительной ясностью и убедительностью показывая, как шла подготовка к Октябрю, какие силы готовились к действию, были введены в действие и действовали в процессе развертывания Великой пролетарской революции, обеспечив победу восставшего пролетариата, поддержанного основными крестьянскими массами. Прочтите — и вы увидите, каким могучим теоретическим оружием владеет партия Ленина — Сталина и как, блестяще владея и действуя этим оружием, партия указала путь победы пролетариату, выдвинувшему вперед закаленные в боях отряды пролетарских революционеров.

«Одна из особенностей Октябрьской революции, — писал товарищ Сталин в 1924 г., — состоит в том, что эта революция является классическим проведением ленинской теории диктатуры пролетариата».[1]

«Вторая особенность Октябрьской революции, — писал Сталин тогда же, — состоит в том, что эта революция является образцом применения на практике ленинской теории пролетарской революции».[2]

Обе эти особенности непосредственно связаны с проблемой отношений между пролетариатом и крестьянством в пролетарской революции. Без правильного построения этих отношений пролетарская революция не могла бы двигаться вперед с таким успехом, с каким это делала Великая Октябрьская революция, преодолевая многочисленные трудности, связанные с решением этой проблемы.

II

ДИКТАТУРА ПРОЛЕТАРИАТА И ПРОЛЕТАРСКАЯ ДЕМОКРАТИЯ

Выше мы уже говорили, что вопрос о победе социализма в одной стране является основным вопросом пролетарской революции. С этой задачей связана судьба пролетарской революции, победоносное развитие которой не может итти иным путем, как путем организации новых, социали-

[1] Сталин. Вопросы ленинизма, стр. 82.
[2] Там же, стр. 86.

стических общественных отношений, организации нового, социалистического общества.

Пролетариат берет в свои руки государственную власть и использует ее именно для того, чтобы, ликвидировав начисто власть буржуазии, опирающейся на частную собственность на средства производства, дающую ей возможность эксплоатировать и угнетать трудящийся народ, освободить труд от эксплоатации капитала, построить общественные отношения на основе обобществления средств производства, использовать прибавочную прибыль в общественных интересах, утвердить социализм.

Основная задача пролетарской революции — установить диктатуру пролетариата, утвердить власть рабочих и крестьян, разрешив противоречия между пролетариатом и крестьянством, использовать государственную власть для построения социалистического общества. Ясно, что отрицание возможности построения социализма в стране, где победила пролетарская революция, логическим последствием должно было бы иметь... ликвидацию победы пролетарской революции.

Борьба за генеральную линию нашей партии, за возможность построения в СССР социализма против троцкистско-зиновьевских отщепенцев, отрицавших возможность построения социализма, была вместе с тем и борьбой за пролетарскую революцию, за право пролетариата не только бороться с буржуазией, но и победить буржуазию.

Известно, что предатели из лагеря зиновьевцев-бухаринцев заявляли, что строить социализм в СССР, конечно, можно, но построить социализм нельзя, так как для этого нет-де необходимых условий.

И. В. Сталин справедливо говорил еще в 1927 г., что такая постановка вопроса означает капитуляцию перед капиталистическими элементами нашего хозяйства.

Однако задача революции заключалась в том, чтобы преодолеть и ликвидировать капиталистические элементы.

Товарищ Сталин справедливо указывал еще в 1927 г., т. е. задолго до того как была разоблачена измена Зиновьева, что внутренняя логика отрицания последним возможности построения в одной стране социализма приводила к выводу: не надо было брать власть в Октябре 1917 г.

Отрицание возможности построения социализма в СССР означало, таким образом, отрицание диктатуры пролетариата, призыв к капитуляции, к отказу от власти, к ликвидации Советского государства.

Между тем, Советское государство и социалистическое строительство все крепли и крепли, наперекор всяким капитулянтским «теориям» и «теорийкам». Известно, что, в то время как троцкистско-бухаринские «теоретики» из кожи лезли вон со своими лженаучными прогнозами, пугая, как это делал, например, Рыков, предсказаниями о провале и крушении советской власти, с одной стороны, и уверяя, с другой стороны, как это делал, например, Бухарин, что кулак обязательно «врастет в социализм», они же и их соучастники не теряли времени, чтобы подорвать дело социалистического строительства и на практике. Рыков добивался того, чтобы вместо сталинского пятилетнего плана была принята его двухлетка, причем мотивом выдвигалось соображение, что лучше и целесообразнее-де иметь два параллельных плана. Товарищ Сталин разоблачил тогда же эту право-троцкистскую махинацию.

«Двухлетний план, — писал Сталин в своей работе «О правом уклоне в ВКП(б)», — выступил у Рыкова на сцену для того, чтобы потом, в ходе практического осуществления пятилетнего плана, противопоставить пятилетке двухлетку, перестроить пятилетку и приспособить ее к двухлетнему плану, сократив и обкарнав ассигнования на дело индустрии». [1]

Карта рыковых и бухариных была бита. Первая пятилетка осталась неприкосновенной и выполнила свою историческую роль, обеспечив невиданный до того успех в деле организации социалистической индустрии, в деле превращения СССР в страну индустриальную, создания экономической базы для уничтожения классов в СССР и для построения социалистического общества.

«Пересесть с обнищалой мужицкой лошади на лошадь крупной машинной индустрии, — вот какую цель преследовала партия, вырабатывая пятилетний план и добиваясь его осуществления», говорил товарищ Сталин в историческом докладе «Итоги первой пятилетки» (1933). Успешному решению этой задачи мешали бухарины, рыковы и их приспешники. Партия и рабочий класс сбросили их в мусорную яму. Партия и рабочий класс выполнили свой долг, решили историческую задачу.

В борьбе против социалистического строительства враги народа открыто выступали против мер, направленных на обуздание кулака,

[1] Сталин. Вопросы ленинизма, стр. 255.

против чрезвычайных мер, доказывая их принципиальную неприемлемость при советской власти.

В своей брошюре «Путь к социализму и рабоче-крестьянский союз» Бухарин откровенно выступил в защиту кулаков, требуя устранения «произвола, хотя бы даже и революционного», требуя отказа от вмешательства власти в ход экономической жизни, даже если такой «произвол» и такое «вмешательство» были бы необходимы в интересах бедняцко-середняцкого крестьянства. Дело, конечно, было не в «произволе», о котором пищали троцкистско-бухаринские «демократы». Дело было в глубоких принципиальных разногласиях, отделявших этих отщепенцев уже в те годы от партии, правильно и решительно действовавшей против кулачества как класса, нашедшего в троцкистско-бухаринской группке своих идеологов и адвокатов. Крики бухаринцев-зиновьевцев о «произволе» и требование «невмешательства» в ход экономической жизни были в те годы подготовкой к более решительной борьбе контрреволюционных группок против мероприятий нашей партии и правительства, направленных на обеспечение победы социализма в нашей стране, к борьбе против программы социалистического переустройства экономических и общественных отношений в СССР.

Важно отметить, что в то же время троцкистско-бухаринское отребье непрочь было «поцарапаться» с кулаком — для отвода глаз, для создания видимости «классовой борьбы» и наступления на кулака.

«В 1926—1927 гг. зиновьевско-троцкистская оппозиция,— как указывал на это товарищ Сталин,— усиленно навязывала партии политику немедленного наступления на кулачество».[1]

Партия не пошла по этому пути, ибо в то время не созрели необходимые для такого серьезного дела условия. На этот путь партия стала лишь в 1929 г., когда были налицо все условия для ликвидации кулачества как класса и замены его производства производством колхозов и совхозов.

Партия не могла бы столь блестяще осуществить план ликвидации кулачества как класса на основе сплошной коллективизации, если бы она не была вооружена подлинно научной марксистско-ленинско-сталинской теорией, с одной стороны, и если бы, с другой стороны, труд-

---

[1] Сталин. Вопросы ленинизма, стр. 291.

щиеся СССР не создали к тому времени сильного и могучего единством советского народа социалистического государства рабочих и крестьян, государства, о котором Ленин говорил, что в период перехода от капитализма к коммунизму, в период свержения буржуазии и полного уничтожения ее, в период невиданно острых форм классовой борьбы, государство «неизбежно должно быть государством *по новому* демократическим (для пролетариев и неимущих вообще) и *по новому* диктаторским (против буржуазии)».[1]

Всячески стараясь расшатать и подорвать диктатуру пролетариата, враги народа не останавливались ни перед какими средствами борьбы. Заговоры, измена, террор, саботаж, вредительство, диверсии — все было пущено в ход, чтобы остановить поступательное движение страны Советов, из года в год укреплявшей свои силы и превращавшейся все более и более в передовую, могучую социалистическую индустриальную державу.

Враги народа правильно учитывали, что основой Советского государства, как государства нового типа, является именно диктатура пролетариата.

В «Государстве и революции», как и в ряде своих позднейших бессмертных научных трудов, Ленин дал законченное изложение марксистского понимания диктатуры пролетариата. Ленин очистил это понимание от оппортунистических извращений каутскианцев и бернштейнианцев, отвергавших, в сущности, диктатуру пролетариата. Ленин указывал, в качестве наиболее яркого доказательства этого, на брошюру Каутского, вышедшую в августе 1918 г., — «Диктатура пролетариата», этот «образец мещанского искажения марксизма и подлого отречения от него *на деле*, при лицемерном признании его *на словах*».[2]

Это «подлое отречение» от марксизма заключается главным образом в том, что каутскианцы, меньшевики, эсеры и другие оппортунисты не доводят признания классовой борьбы до периода п е р е х о д а от капитализма к коммунизму, а это, как указывает Ленин, и есть самое главное в учении о классовой борьбе.

---

[1] Л е н и н. Соч., т. XXI, стр. 393.
[2] Там же, стр. 392.

Заслуга Ленина не только в том, что он очистил марксизм от оппортунистических искажений и мещанской пошлятины Каутского и К°. Заслуга Ленина также в том, что он разработал всю *систему* пролетарской диктатуры, показал все исключительное значение диктатуры пролетариата и ее особенности в условиях советской власти, а также черты, отличающие ее от диктатуры буржуазии.[1]

Отметим лишь наиболее важные элементы этой системы, ее основные положения.

Ленин рассматривает диктатуру пролетариата, раньше всего, как «*продолжение* классовой борьбы пролетариата, в *новых* формах», а пролетарское государство — как «*орудие* пролетариата в его классовой борьбе. Особая *дубинка*, — поясняет Ленин, — rien de plus!» (ничего больше).[2]

Эта классовая борьба при диктатуре пролетариата приобретает новые формы соответственно новым задачам; таких «*пять* новых (главнейших) задач и respective (соответственно) новых форм», а именно: а) «*подавление сопротивления эксплуататоров*»; ... б) «*гражданская война*»; в) «*нейтрализация*» *мелкой буржуазии, особенно крестьянства*»; г) «*использование*» *буржуазии*» (вопрос об использовании специалистов); д) «*воспитание новой дисциплины*».[3]

Ленин рассматривает диктатуру пролетариата как разрушение буржуазной демократии и создание пролетарской демократии, подчеркивая фальшь и лицемерие первой и реальность второй демократии. Ленин, например, так разоблачает буржуазное равенство: «Все равны, *не считая* денег, капитала, земли...», или: «Империалистская война 1914—1918, как «последнее слово» буржуазной демократии», или: «Диктатура буржуазии, прикрытая парламентскими формами».[4]

Далее Ленин останавливается на теме «Диктатура пролетариата и особенности империализма». Здесь Ленин рекомендует читателю ознакомиться с трудом «Империализм, как новейший этап капитализма», где он указывает на те работы Маркса и Энгельса, в которых говорится

---

[1] См. Ленин. Соч., т. XXV, стр. 5—12.
[2] См. там же, стр. 5.
[3] См. там же, стр. 6—7.
[4] См. там же, стр. 8—9.

о продажности «рабочих приказчиков класса капиталистов». Ленин добавляет: «2 главных «струи»: продажные и филистеры» и заканчивает тезисом: «2 *Интернационала.* Диктатура *революционных* элементов класса. Одна страна и весь мир».[1] (Имеется в виду проблема захвата власти пролетариатом в одной стране.)

В четвертом разделе, говоря о диктатуре пролетариата и советской власти, Ленин бичует «невежество и тупость *вождей* II Интернационала», признававших «Советы для борьбы, *но не* для государственной власти!!». Ленин констатирует «триумфальное шествие советской идеи по всему миру». «Открыта (массовым движением пролетариата) форма диктатуры пролетариата!!». «III Интернационал». «Советская конституция Р.С.Ф.С.Р. § *23* ее[2] *N. B.*»[3]

Так Ленин определил сущность пролетарской диктатуры и ее роль в пролетарской революции и в пролетарском государстве.

Сталин в своей работе «Об основах ленинизма» писал о диктатуре пролетариата: «Победа диктатуры пролетариата означает подавление буржуазии, слом буржуазной государственной машины, замену буржуазной демократии демократией пролетарской».[4]

Сталин указал при этом на то важное обстоятельство, что новым задачам диктатуры пролетариата должны соответствовать и новые формы организации пролетариата, так как старые формы ее, выросшие на основе буржуазного парламентаризма, оказываются уже недостаточными.

«Этой новой формой организации пролетариата являются Советы».[5]

Советы, советская власть являются наиболее демократической формой государственной власти, ибо Советы — непосредственные организации народных масс, наиболее мощные органы революционной борьбы масс, всеобъемлющие массовые организации пролетариата, крестьян, всех трудящихся; Советы — органы государственной власти, обеспечивающие постоянное, непременное, решающее участие народа в демократическом управлении государством.

---

[1] Ленин. Соч., т. XXV, стр. 11.
[2] § 23 Конституции лишает буржуазию избирательного права.
[3] Ленин. Соч., т. XXV, стр. 11—12.
[4] Сталин. Вопросы ленинизма, стр. 32.
[5] Там же.

2*

Диктатура пролетариата есть классовый союз пролетариата и трудящихся масс крестьянства для свержения капитала, для окончательной победы социализма, при условии, что руководящей силой этого союза является пролетариат.

Таково определение диктатуры пролетариата, данное товарищем Сталиным. Такова сущность диктатуры пролетариата, не только не противоречащая интересам трудящегося крестьянства, но полностью отвечающая его интересам.

Именно такое понимание диктатуры пролетариата обеспечивает направление всей политики Советского государства, которая ведет к упрочению и расцвету советской страны, опирающейся на нерушимый союз рабочих и крестьян.

Ленин указывает, что «научное понятие диктатуры означает не что иное, как ничем не ограниченную, никакими законами, никакими абсолютно правилами не стесненную, непосредственно на насилие опирающуюся власть».[1]

Буржуазия и ее агенты хотят это положение толковать как оправдание произвола и беззакония. Но такие попытки не давали до сих пор результата, не дадут никакого результата и впредь, ввиду явной их несостоятельности.[2]

Советское государство, опирающееся на диктатуру пролетариата, не связано законами другого класса, законами свергнутой власти, построенной другим классом. Советское государство регулирует общественные отношения при помощи *советских* законов, сущность которых — в обеспечении интересов трудящихся при помощи принудительных мер, насилия, но не только насилия.

«Диктатура, — говорит Ленин, — означает не только насилие, хотя она невозможна без насилия, она означает также организацию труда более высокую, чем предыдущая организация».[3]

Дело в том еще, говорит Ленин, что «пролетариат представляет

---

[1] *Ленин.* Соч., т. XXV, стр. 441.

[2] Достаточно указать на многочисленные высказывания Ленина и Сталина о важности точного выполнения советских законов и соблюдения дисциплины, установленной советскими законами. См., например, письмо Ленина по случаю победы над Колчаком, «О двойном подчинении и законности» и ряд других работ.

[3] *Ленин.* Соч., т. XXIV, стр. 305.

и осуществляет более высокий тип общественной организации труда, по сравнению с капитализмом. В этом суть. В этом источник силы и залог неизбежной полной победы коммунизма».[1]

Известна классическая формула Сталина о трех основных сторонах диктатуры пролетариата:

«1) Использование власти пролетариата для подавления эксплоататоров, для обороны страны, для упрочения связей с пролетариями других стран, для развития и победы революции во всех странах.

2) Использование власти пролетариата для окончательного отрыва трудящихся и эксплоатируемых масс от буржуазии, для упрочения союза пролетариата с этими массами, для вовлечения этих масс в дело социалистического строительства, для государственного руководства этими массами со стороны пролетариата.

3) Использование власти пролетариата для организации социализма, для уничтожения классов, для перехода в общество без классов, в общество без государства».[2]

Сталин дальше развил учение о диктатуре пролетариата, показав, что диктатура пролетариата не представляет собой чего-либо застывшего и неподвижного, данного раз навсегда, действующего независимо от исторических условий и обстановки. Нет, по Сталину — «диктатура пролетариата имеет свои периоды, свои особые формы, разнообразные методы работы».[3]

Одни методы диктатура пролетариата применяет в период гражданской войны, когда «особенно бьет в глаза насильственная сторона диктатуры», что, однако, не означает отсутствия строительной работы. Другие методы действуют в период строительства социализма, когда, «наоборот, особенно бьет в глаза мирная, организаторская, культурная работа диктатуры, революционная законность и т. д.».[4]

И в тот и в другой периоды необходимы органы подавления, армия и другие организации, так как без них невозможна сколько-нибудь обеспеченная строительная работа диктатуры.

Сталин показал также «механизм» диктатуры пролетариата: «привода»,

---

[1] Ленин. Соч., т. XXIV, стр. 336.

[2] Сталин. Вопросы ленинизма, стр. 117.

[3] Там же.

[4] Там же.

«рычаги», «направляющую силу», совокупность которых составляет систему диктатуры пролетариата, о которой говорил Ленин.

Здесь мы видим: *партию* (основную руководящую силу в системе диктатуры пролетариата), высшую форму классового объединения пролетариата (Ленин); *профсоюзы* (школа коммунизма; они соединяют рабочие массы с авангардом рабочего класса); *Советы* (они соединяют миллионные массы трудящихся с авангардом пролетариата, они — прямое выражение диктатуры пролетариата); *кооперацию* (облегчает связь авангарда с массами крестьянства и создает возможность вовлечения последних в русло социалистического строительства); *союз молодежи* (помогает партии, дает резервы для всех остальных массовых организаций по всем отраслям управления).[1]

Не ясно ли, что диктатура пролетариата — это инструмент массового порядка, орган, выражающий волю народных масс, сливающихся со своим авангардом, пользующимся безграничной поддержкой и доверием пролетарской, в первую голову, но также и непролетарской части трудящихся.

При таких условиях дело оказывается не в насилии, хотя без насилия не может обойтись диктатура пролетариата, а в доверии, в братской связи рабочего класса и его партии со всеми трудящимися.

Авангард пролетариата должен быть учителем, руководителем, вождем всех трудящихся и эксплоатируемых в деле устройства своей общественной жизни без буржуазии и против буржуазии.

Так учат Ленин и Сталин. Такую роль и играет в жизни Советского Союза авангард рабочего класса — Коммунистическая партия большевиков, организованная, выпестованная и подготовленная Лениным и Сталиным к своей великой исторической роли.

☆

Крестьянский вопрос в СССР — это вопрос о взаимоотношении пролетариата и крестьянства. Это важнейший вопрос, особенно если принять во внимание удельный вес крестьянства в СССР. Не случайно взоры всех врагов социализма и социалистической революции в СССР были обра-

---

[1] См. Сталин. Вопросы ленинизма, стр. 119—120.

щены в эту сторону, в надежде, что большевики и руководимый ими пролетариат не сумеют найти правильное решение этого вопроса, сорвутся на крестьянском вопросе. Не случайно главными героями буржуазной контрреволюции оказались наряду с генералами Деникиным, Колчаком, Врангелем и Юденичем вожаки эсеров — Чернов, Керенский, Брешковская, Савинков и т. п. Не случайно и то, что троцкисты и зиновьевцы били в эту же точку, стараясь своими провокаторскими рецептами «завинчивания гайки» в отношении крестьянства взорвать рабоче-крестьянский союз.

Мудрая ленинско-сталинская политика в крестьянском вопросе устранила эту опасность. Ленинско-сталинская наука строительства социализма обеспечила прочность и нерушимость этого союза, явившегося фундаментом всего дела социализма в нашей стране.

Ленин так определял политику пролетарского государства в отношении крестьянства:

«Задача здесь (в хозяйственных отношениях со средним крестьянством. — *А. В.*), — говорил Ленин, — сводится не к экспроприации среднего крестьянина, а к тому, чтобы учесть особенные условия жизни крестьянина, к тому, чтобы учиться у крестьян способам перехода к лучшему строю и *не сметь командовать!* Вот правило, которое мы себе поставили».[1]

Ленин подчеркивал, что «миллионы не могут сразу понять перемену курса» и что «новые условия и новые задачи по отношению к этому классу требуют новой психологии».[2]

Основа этой психологии — в новом понимании положения и роли среднего крестьянства в пролетарской революции. Главное здесь — правильно поставить задачу, как говорил Ленин, устроения жизни среднего крестьянства.

«Мы должны с ним жить в мире, — учит Ленин. — Среднее крестьянство в коммунистическом обществе только тогда будет на нашей стороне, когда мы облегчим и улучшим экономические условия его жизни. Если бы мы могли дать завтра 100 тысяч первоклассных тракторов, снабдить их бензином, снабдить их машинистами (вы прекрасно знаете, что пока

[1] Ленин. Соч., т. XXIV, стр. 168.
[2] Там же, стр. 168—169.

это — фантазия), то средний крестьянин сказал бы: «Я за коммунию» (т. е. за коммунизм)».[1]

Ленин и Сталин учат рассматривать трудящиеся массы крестьянства как резерв пролетариата, как прочную опору в проведении широчайших и коренных мероприятий по организации социалистических отношений.

Сталин напоминает слова Энгельса, который в 90-х годах прошлого столетия в своей книге «Крестьянский вопрос во Франции и Германии» писал: социалистическая партия, «чтобы завоевать политическую власть, ... должна сначала из города пойти в деревню, должна сделаться силой в деревне».[2] Но что значит «пойти из города в деревню?» Как «сделаться силой в деревне?»

На эти вопросы Энгельс отвечал, что нужно делать все возможное, чтобы крестьянину лучше жилось и чтобы облегчить ему переход к товариществу, если он на это решится; нужно, если он не решится на это, дать крестьянину возможность подумать об этом на своем клочке; нужно оказать крестьянам материальную поддержку из общественных средств, будучи очень щедрыми в этом деле, так как затраты окупятся впоследствии во много раз при общественном преобразовании в целом.

И. В. Сталин, комментируя высказывания Энгельса, подчеркивал, что именно в стране диктатуры пролетариата указанное Энгельсом может быть осуществлено наиболее легко и полно, тем более что такие меры уже тогда проводились в жизнь в СССР.

«Как можно отрицать, — говорил товарищ Сталин, — что это обстоятельство, в свою очередь, должно облегчить и двинуть вперед дело хозяйственного строительства Советской страны?»[3]

Все последующие события полностью подтвердили это замечание товарища Сталина, показав, что успехи хозяйственного и вообще государственного строительства органически связаны с участием в этом строительстве трудящегося крестьянства, первого и верного помощника рабочего класса в борьбе за социализм.

Диктатура пролетариата явилась прочным фундаментом братского сотрудничества рабочего класса и основных масс крестьянства. Вся прак-

---

[1] Ленин. Соч., т. XXIV, стр. 170.
[2] Маркс и Энгельс. Соч., т. XVI, ч. 2, стр. 442.
[3] Сталин. Вопросы ленинизма, стр. 43.

тика советской власти служит живым и полным опровержением троцкистско-бухаринского «тезиса» о якобы коренных противоречиях между рабочим классом и трудящимся крестьянством.

Братское сотрудничество рабочего класса и крестьянства — прямое следствие советской системы, покоящейся на союзе рабочих и крестьян; она всеми своими особенностями и качествами сближает и объединяет, а не разъединяет рабочих и крестьян.

Ленин, определяя сущность советской власти, писал о шести особенностях Советов. Одну из этих особенностей Ленин видел в том, что новый государственный аппарат, каковым являются Советы, «дает форму организации авангарда, т.-е. самой сознательной, самой энергичной, передовой части *угнетенных* классов, рабочих и крестьян, являясь таким образом аппаратом, посредством которого авангард угнетенных классов может поднимать, воспитать, обучать и вести за собой *всю гигантскую массу* этих классов, до сих пор стоявшую совершенно вне политической жизни, вне истории».[1]

Особенностью пролетарской революции является то, что она в состоянии и должна связать многомиллионные массы трудящихся с пролетариатом в длительный союз. Это же является и особенностью Советского государства как особой формы союза рабочих и крестьян.

Об этой особой форме союза товарищ Сталин в своем труде «К вопросам ленинизма» писал так: «Состоит она, эта особая форма союза, в том, что руководящей силой этого союза является пролетариат. Состоит она, эта особая форма союза, в том, что руководителем государства, руководителем в системе диктатуры пролетариата является *одна* партия, партия пролетариата, партия коммунистов, которая *не делит и не может делить* руководства с другими партиями».[2]

«В этом суть. В этом источник силы и залог неизбежной полной победы коммунизма».[3]

Для победы коммунизма нужно сделать «громадный шаг вперед в развитии производительных сил, надо преодолеть сопротивление (часто пассивное, которое особенно упорно и особенно трудно поддается преодо-

---

[1] Ленин. Соч., т. XXI, стр. 258.
[2] Сталин. Вопросы ленинизма, стр. 115.
[3] Ленин. Соч., т. XXIV, стр. 336.

лению) многочисленных остатков мелкого производства, надо преодолеть громадную силу привычки и косности, связанной с этими остатками».[1]

Но только пролетариат способен преодолеть эту громадную силу пережитков капитализма в экономике и психологии людей, несмотря на то, что сам пролетариат в первый период революции не свободен еще от этих пережитков.

Ведущая роль пролетариата в переустройстве старых общественных отношений вырастает из самого характера этих отношений. Она определяется общественным положением пролетариата в производстве, а это главенствующий фактор общественных отношений. В переходный период от капитализма к социализму пролетариат неизбежно сохраняет положение ведущего класса в обществе со всеми вытекающими отсюда последствиями. Это один из важнейших законов пролетарской революции.

Самое понятие диктатуры пролетариата означает руководство со стороны пролетариата непролетарскими трудящимися массами. Именно так определяют Ленин и Сталин понятие диктатуры пролетариата.

Ленинизм учит, что только пролетариат в состоянии руководить всей массой трудящихся в борьбе за их освобождение, в борьбе за удержание и укрепление победы.

В работе Ленина «Великий почин» эта особенность пролетарской диктатуры подчеркивается в следующих словах:

«Чтобы победить, чтобы создать и упрочить социализм, пролетариат должен решить двоякую или двуединую задачу: во-первых, увлечь своим беззаветным героизмом революционной борьбы против капитала всю массу трудящихся и эксплуатируемых, увлечь ее, организовать ее, руководить ею для свержения буржуазии и полного подавления всякого с ее стороны сопротивления; во-вторых, повести за собой всю массу трудящихся и эксплуатируемых, а также все мелко-буржуазные слои, ... на путь создания новой общественной связи, новой трудовой дисциплины, новой организации труда, соединяющей последнее слово науки и капиталистической техники с массовым объединением сознательных работников, творящих крупное социалистическое производство.

Эта вторая задача труднее первой, ибо она ни в коем случае не может быть решена героизмом отдельного порыва, а требует самого длительного, самого упорного, самого трудного героизма массовой и *будничной*

---

[1] Ленин. Соч., т. XXIV, стр. 337.

работы. Но эта задача и более существенна, чем первая, ибо в последнем счете самым глубоким источником силы для побед над буржуазией и единственным залогом прочности и неотъемлемости этих побед может быть только новый, более высокий способ общественного производства, замена капиталистического и мелко-буржуазного производства крупным социалистическим производством».[1]

Союз рабочих и крестьян под руководством рабочего класса и при руководящем положении в СССР партии коммунистов — вот основа Советского государства, новой по своему типу формы государственной власти, советской власти. Опыт революции доказал, что эта форма власти — гораздо более высокая форма демократии, чем какая бы то ни было другая, известная истории. Это понял и знает советский народ, ощущающий на каждом шагу хозяйственного и культурного строительства в СССР все преимущества советской власти по сравнению с другою властью.

Нет необходимости — да в настоящей статье и нет достаточной для этого возможности — давать обзор тех конкретных мероприятий, которые показали бы, как на практике претворялась политика Советского государства в отношении крестьянства. Путь от декрета о земле и наделения крестьян землей, которую крестьянство получило от советской власти, от комбедов, явившихся боевыми органами крестьянской бедноты в борьбе против кулачества, до массового колхозного движения и массовой организации колхозов, решивших судьбу кулачества как класса, — у всех на памяти. Великий путь, указанный беднейшим и середняцким массам крестьянства партией Ленина — Сталина, — это путь спасения, путь к зажиточной жизни миллионов пахарей, трудящихся на родной земле, путь счастливого труда в союзе и под руководством рабочего класса, верного друга и бесстрашного защитника интересов всех трудящихся. Под водительством рабочего класса и его партии трудящиеся СССР построили свое социалистическое государство, несокрушимый оплот их свободы, независимости и счастья...

☆

Советское государство, как государство нового типа, ново своим демократизмом — последовательным социалистическим демократизмом.

---

[1] Ленин. Соч., т. XXIV, стр. 339.

Советский демократизм обеспечивает господство в обществе воли большинства рабочих и крестьян, большинства трудящихся, участвующих в управлении государством. Именно советская форма демократизма пробуждает к активной жизни, к социалистическому творчеству миллионы людей, обеспечивает выделение из народных масс тысяч и тысяч руководителей и организаторов. В качестве одной из важнейших особенностей советской власти Ленин отмечал то обстоятельство, что именно советская власть, формы и методы работы органов Советского государства обеспечивают решение задачи выделения и воспитания новых руководителей и организаторов из народа.

Тысячи и тысячи раз подтверждаются слова Ленина и Сталина о советском демократизме как высшей форме демократии, на деле обеспечивающей за миллионными народными массами участие в государственном управлении, в строительстве социализма. Наиболее ярким свидетельством последовательного, до конца развернутого социалистического демократизма явилась Великая Сталинская Конституция.

Советская власть является формой наиболее полного и развернутого демократизма. Она является в то же время и формой диктатуры рабочего класса, обеспечивающей самую возможность демократии для народа. Советская демократия и пролетарская диктатура суть две стороны одного и того же явления.

Советское государство сильно инициативой народных масс, их творчеством и беззаветной преданностью делу своего освобождения, своей свободы, делу славы и возвеличения своей социалистической родины. Вот откуда вырастает животворный патриотизм — источник великих побед, великих подвигов советских людей.

В первые дни советской власти Ленин учил необходимости применять максимально разнообразные методы и приемы в деле организации новых общественных отношений в Республике Советов. Говоря о задаче перехода «к крупному, на машинной индустрии построенному хозяйству», к социализму, Ленин писал, что конкретные условия и формы этого перехода неизбежно являются и должны быть разнообразными в зависимости от тех условий, при которых начинается движение, направленное к созданию социализма.

Ленин писал: «Чем больше будет такого разнообразия, — конечно, если оно не перейдет в оригинальничание, тем вернее и тем быстрее будет

обеспечено, как достижение нами демократического централизма, так и осуществление социалистического хозяйства. Нам остается теперь только организовать соревнование. ...Придавленные капиталистическим строем, мы не можем в точности даже представить себе в настоящее время, какие богатые силы таятся в массе трудящихся, в разнообразии трудовых коммун большого государства, в интеллигентских силах, которые до сих пор работали, как мертвые безгласные исполнители предначертаний капиталистов, какие силы таятся и могут развернуться при социалистическом устройстве общества. Наше дело состоит только в том, чтобы расчистить дорогу всем этим силам. И, если мы поставим организацию соревнования, как нашу государственную задачу, — то, при условии проведения советских принципов государственного порядка, при условии уничтожения частной собственности на землю, на фабрики, на заводы и пр., — результаты должны будут показать себя неизбежно и подскажут нам дальнейшие формы строительства».[1]

В статье «Как организовать соревнование?» Ленин писал, что при капитализме задушена инициатива масс, так как при капитализме конкуренция «означает неслыханно зверское подавление предприимчивости, энергии, смелого почина *массы* населения, гигантского большинства его, девяносто девяти сотых трудящихся, означает также замену соревнования финансовым мошенничеством, деспотизмом, прислужничеством на верху социальной лестницы».[2]

«Социализм, — пишет далее Ленин, — не только не угашает соревнования, а напротив, впервые создает возможность применить его действительно *широко*, действительно в *массовом* размере, втянуть действительно большинство трудящихся на арену такой работы, где они могут проявить себя, развернуть свои способности, обнаружить таланты, которых в народе — непочатой родник и которые капитализм мял, давил, душил тысячами и миллионами».[3]

Ленин звал народные массы к творчеству и инициативе. Он указывал, что в процессе соревнования «должны *практически* выделяться и выдвигаться наверх, в дело общегосударственного управления, *организатор-*

---

[1] Ленин. Соч., т. XXII, стр. 416—417.
[2] Там же, стр. 158.
[3] Там же.

ские таланты. Их много в народе»,[1] говорил Ленин. За годы советской власти выдвинулись тысячи и тысячи талантливых, вышедших из самой гущи народа организаторов и руководителей во всех областях государственного строительства: инженеры, директоры, полководцы и строители Советской Армии, деятели науки, выдающиеся государственные работники, врачи, юристы, артисты, писатели, художники — мастера самых разнообразных видов деятельности.

Тысячи и десятки тысяч героев и героинь труда и боевого фронта родила и воспитала наша отчизна, никогда не видевшая до советского строя такого обилия замечательных людей — творцов и строителей новой жизни.

Ленин предвосхитил нарождение новых форм социалистического соревнования, которые, как, например, стахановское движение, дали громадный толчок росту производительности труда и общему подъему хозяйственного социалистического строительства. Впоследствии, говоря о стахановском движении, товарищ Сталин справедливо отмечал, что «оно подготовляет условия для перехода от социализма к коммунизму».[2]

☆

Вопрос о государстве, естественно, приобретает в условиях пролетарской революции особенно важное значение, ибо это был вопрос о власти, а вопрос о власти, как учит Ленин, является основным вопросом революции.

Маркс и Энгельс дали замечательное, систематическое изложение своих взглядов о государстве, показав решающее значение государства для пролетарской революции и для организации социалистического общества.

Классическая формула «Коммунистического манифеста», написанного Марксом и Энгельсом, гласит, что «первым шагом в рабочей революции является превращение пролетариата в господствующий класс, завоевание демократии.

Пролетариат использует свое политическое господство для того, чтобы

---

[1] Ленин. Соч., т. XXII, стр. 167.
[2] Сталин. Вопросы ленинизма, стр. 495.

вырвать у буржуазии шаг за шагом весь капитал, централизовать все орудия производства в руках государства, т. е. пролетариата, организованного, как господствующий класс, и возможно более быстро увеличить сумму производительных сил».[1]

Марксизм исходит, таким образом, из необходимости для пролетариата овладеть государством, стать «господствующим классом» и использовать государство для централизации в его руках всех орудий производства.

Марксизм исходит также из того, что превращение капиталистического общества в коммунистическое займет целый период, что это превращение осуществится революционным путем и что государство этого переходного политического периода и не может быть не чем иным, кроме как революционной диктатурой пролетариата.

Диктатура пролетариата выполняет очистительную миссию, ломая сопротивление эксплоататоров и обеспечивая громадное расширение демократии, становящейся впервые демократией для бедных — «демократизмом для народа, а не демократизмом для богатеньких» (Ленин). Диктатура пролетариата неизбежно устанавливает изъятия из свободы по отношению к эксплоататорам.

Ленин подчеркивал, что это прекрасно выразил в письме к Бебелю Энгельс, сказав, что «пролетариат нуждается в государстве не в интересах свободы, а в интересах подавления своих противников, а когда можно будет говорить о свободе, не будет государства».[2]

Учение о государстве — важнейшая часть марксистского учения о революции, марксистской социалистической теории. Значение теории Маркса превосходно определил Ленин в 90-х годах прошлого столетия, когда в статье «Наша программа» сформулировал задачи революционно-социалистической партии. Ленин писал, что теория Маркса выяснила настоящую задачу этой партии: «не сочинение планов переустройства общества, не проповедь капиталистам и их прихвостням об улучшении положения рабочих, не устройство заговоров, *а организацию классовой борьбы пролетариата и руководство этой борьбой, конечная цель*

---

[1] К. Маркс и Ф. Энгельс. Манифест Коммунистической партии, М., 1939, стр. 47.

[2] Ленин. Соч., т. XXI, стр. 431.

которой — завоевание политической власти пролетариатом и организация социалистического общества».[1]

Объявляя себя всецело сторонником теории Маркса, Ленин подчеркивал в то же время, что «мы вовсе не смотрим на теорию Маркса как на нечто законченное и неприкосновенное».

Ленин указывал, что теория Маркса положила лишь «краеугольные камни той науки, которую социалисты *должны* двигать дальше во всех направлениях, если они не хотят отстать от жизни».

«Мы думаем, что для русских социалистов особенно необходима *самостоятельная* разработка теории Маркса, ибо эта теория дает лишь общие *руководящие* положения, которые применяются *в частности* к Англии иначе, чем к Франции, к Франции иначе, чем к Германии, к Германии иначе, чем к России».[2]

Эти указания Ленина носят глубоко принципиальный характер и чрезвычайно важны для правильного понимания последующего развития марксистской теории, которая вовсе не является «догмой» и развитие которой насущно необходимо в интересах поступательного коммунистического движения во всем мире.

Ленин говорил в другом случае («Проект программы нашей партии»), по поводу Эрфуртской программы, что мы не должны отказываться от подражания этой программе, но это ни в коем случае не должно вести «к забвению *особенностей* России, которые должны найти *полное выражение* в особенностях нашей программы».[3]

Руководствуясь указанными выше принципами, Ленин развил дальше марксистское учение о пролетарской диктатуре и демократии, о государстве и о социалистическом государстве в особенности.

Как известно, главным и основным в учении марксизма о государстве является та новая, по сравнению с «Коммунистическим манифестом», постановка вопроса об отношении пролетариата к буржуазному государству, которую дал Маркс в «18 брюмера Луи Бонапарта».

Это новое — Ленин назвал его громадным шагом вперед по сравнению с «Коммунистическим манифестом» — заключается в «чрезвычайно

---

[1] Ленин. Соч., т. II, стр. 491.
[2] Там же, стр. 492.
[3] Там же, стр. 514.

точном, определенном, практически-осязательном» выводе, к которому в 1852 г. пришел Маркс, а именно: «... все прежние революции усовершенствовали государственную машину, а ее надо разбить, сломать».[1]

Этот вывод — прямой результат «подытожения» (Ленин) Марксом того исторического опыта, который проделала за 1848—1851 гг. Европа. Однако этого опыта оказалось недостаточно для того, чтобы определить, чем заменить эту машину, подлежащую слому и уничтожению.

Показав на анализе империализма силу концентрации банкового капитала с необыкновенным усилением государственной капиталистической машины, с неслыханным ростом ее чиновничьего и военного аппарата, Ленин заключает: «Всемирная история подводит теперь, несомненно, в несравненно более широком масштабе, чем в 1852 году, к «концентрации всех сил» пролетарской революции на «разрушение» государственной машины.

Чем заменит ее пролетариат, об этом поучительнейший материал дала Парижская Коммуна».[2]

Так писал Ленин в августе 1917 г.

Парижская Коммуна, подтвердив, что «рабочий класс не может просто овладеть готовой государственной машиной и пустить ее в ход для своих собственных целей»,[3] показала вместе с тем, что на смену разбитой государственной машине буржуазии придет уже «не парламентское, а работающее учреждение», «в одно и то же время и законодательствующее и исполняющее законы». Значит, вместо парламентаризма с его разделением властей диктатура пролетариата создает новую государственную систему, опирающуюся на единство воли революционного народа.

Ленин так комментирует в «Государстве и революции» это коренное изменение в государственном строе Коммуны:

«Продажный и прогнивший парламентаризм буржуазного общества Коммуна заменяет учреждениями, в коих свобода суждения и обсуждения не вырождается в обман, ибо парламентарии должны сами работать, сами исполнять свои законы, сами проверять то, что получается в жизни,

---

[1] Ленин. Соч., т. XXI, стр. 388.

[2] Там же, стр. 391.

[3] К. Маркс и Ф. Энгельс. Манифест Коммунистической партии. Предисловие к немецкому изданию 1872 г. М., 1939, стр. 10.

сами отвечать непосредственно перед своими избирателями. Представительные учреждения остаются, но парламентаризма, как особой системы, как разделения труда законодательного и исполнительного, как привилегированного положения для депутатов, здесь *нет*. Без представительных учреждений мы не можем себе представить демократии, даже и пролетарской демократии; без парламентаризма можем и *должны*, если критика буржуазного общества для нас не пустые слова, если стремление свергнуть господство буржуазии есть наше серьезное и искреннее стремление, а не «избирательная» фраза для уловления голосов рабочих, как у меньшевиков и эс-эров, как у Шейдеманов и Легинов, Самба и Вандервельдов».[1]

Парижская Коммуна явилась «открытой наконец» пролетарской революцией, несмотря на ее неудачу, государственной формой, при которой может произойти экономическое освобождение труда.

«Коммуна, — писал Ленин, — первая попытка пролетарской революции *разбить* буржуазную государственную машину и «открытая наконец» политическая форма, которою можно и должно *заменить* разбитое.

Мы увидим в дальнейшем изложении, — добавлял Ленин, — что русские революции 1905 и 1917 годов, в иной обстановке, при иных условиях, продолжают дело Коммуны и подтверждают гениальный исторический анализ Маркса».[2]

Мы скажем: гений Ленина подытожил исторический опыт пролетарской классовой борьбы за время, протекшее после Коммуны, и особенно опыт революции 1905 г., и ответил на поставленный почти полвека назад вопрос о той государственной форме, в которой пролетариат осуществит свое действительное освобождение. Этот ответ явился новым гигантским шагом вперед в практике рабочего социалистического движения, в марксистской теории пролетарской революции и государства. Этот ответ гласил: «Для перехода от буржуазного строя к социалистическому, для диктатуры пролетариата, республика Советов рабочих, солдатских и крестьянских депутатов является не только формой более высокого типа демократических учреждений (по сравнению с обычной буржуазной республи-

[1] Ленин. Соч., т. XXI, стр. 402.
[2] Там же, стр. 408.

кой при Учредительном собрании, как венце ее), но и единственной формой, способной обеспечить наиболее безболезненный переход к социализму».[1]

Советская власть явилась «той искомой и найденной, наконец, политической формой, в рамках которой должно быть совершено экономическое освобождение пролетариата, полная победа социализма».[2]

Заслуга открытия такой политической формы принадлежит Ленину. Заслуга укрепления и победоносного развития этой формы принадлежит великому продолжателю дела Ленина — нашему учителю и вождю Сталину.

III

СОВЕТСКОЕ ГОСУДАРСТВО И НАЦИОНАЛЬНЫЙ ВОПРОС

Национальный вопрос является одним из наиболее трудных и острых вопросов внутренней политики, особенно в государствах с многонациональным составом. В странах буржуазной демократии, даже в наиболее передовых, как США, Франция, Великобритания, национальный вопрос не решен и не может быть решен в интересах национальных меньшинств, поскольку вся система капиталистических общественных отношений не благоприятствует такому решению. Капиталистическая эксплоатация в области экономических отношений накладывает свой тлетворный отпечаток и на общественные отношения во всех других областях жизни. Беспощадная эксплоатация слабого и отчаянная конкуренция «свободных» товаровладельцев на «свободном» рынке порождают неизбежность использования расовых, национальных и религиозных предрассудков в целях обеспечения наиболее сильным экономически и политически группам привилегированного положения в борьбе за рынок. Это определяет соответствующие политические и правовые последствия, с одной стороны, и последствия морально-бытового порядка — с другой.

Негритянский вопрос в США, индусский или ирландский вопрос в Великобритании, еврейский вопрос во Франции — все эти «вопросы»

---

[1] Ленин. Соч., т. XXII, стр. 131. Тезисы об Учредительном собрании, тезис 3.

[2] Сталин. Вопросы ленинизма, стр. 35.

3*

являются порождением капитализма, органически связаны с капитализмом и не могут быть правильно решены в рамках капитализма. Разумеется, формально, юридически ни одна малая нация в этих странах не поставлена, как правило, в худшее положение, чем господствующее большинство. Но на деле таких ограничений и дискриминаций не мало. Не случайно же в американский парламент за последние 50 лет не удалось быть избранным ни одному негру. Между тем американская конституция (1787 г.) провозглашает равенство избирательных прав всех американских граждан, а закон 1870 г. торжественно заявляет, что участие в выборах «не будет отрицаться и ограничиваться Соединенными Штатами или отдельными штатами под предлогом расы, цвета кожи или прежнего рабского состояния».[1]

В стране социализма, в социалистическом государстве рабочих и крестьян, нет места национальному вопросу. Такого вопроса в СССР нет и не может быть, ибо нет той питающей этот вопрос среды, которая характерна для капиталистических стран и отсутствие которой характерно для социалистической страны.

Дело Дрейфуса в республиканской Франции — разве не легло оно несмываемым пятном позора на буржуазную демократию, даже спустя сто с лишним лет после провозглашения буржуазной Республики не разделавшуюся окончательно с антисемитизмом и националистическим шовинизмом!

В СССР национальный вопрос решен на основе принципов ленинско-сталинской национальной политики, последовательно и неуклонно проводящейся в жизнь с первых дней образования Советского государства.

Первые декреты советской власти, как вехи гигантского переустройства всех общественных отношений — хозяйственных и политических, обозначали путь развития советской государственной системы, строившейся и построенной на новых началах, во имя осуществления новых социальных целей и идеалов.

Национальный вопрос явился одним из коренных вопросов, решение которого в значительной степени определило все содержание великих преобразовательных мероприятий советской государственной власти,

---

[1] См. «Конституции буржуазных государств», т. I, стр. 19, изд. 1935 г.

решение которого обусловило целый ряд важнейших особенностей советской внутренней и внешней политики.

Закладывая основы нового советского общественного и государственного строя, советская власть с первых же шагов своей деятельности обратила особое внимание на разрешение национального вопроса, имевшего исключительное значение в нашей стране, отличавшейся таким многонациональным составом населения.

Особым значением национального вопроса в России объясняется и тот факт, что впервые в истории государств в Советском государстве был создан Народный комиссариат по национальным делам, во главе с товарищем Сталиным.

Первым актом советского правительства в области национальных дел явилась «Декларация прав народов России», написанная товарищем Сталиным в первые же дни Октябрьской революции.

«Декларация» провозглашала политику «добровольного и честного союза народов России». Подтверждая уже ранее (II Съезд Советов) провозглашенное право народов на самоопределение, Совет Народных Комиссаров так определял основы советской национальной политики:

«1) Равенство и суверенность народов России.

2) Право народов России на самоопределение вплоть до отделения и образования самостоятельного государства.

3) Отмена всех и всяких национальных и национально-религиозных привилегий и ограничений.

4) Свободное развитие национальных меньшинств и этнографических групп, населяющих территорию России».[1]

Практическое применение принципов «Декларации» нашло свое выражение в декретах о независимости Украинской Республики, о Финляндской Республике, в декрете и обращении товарища Сталина «О Турецкой Республике», положивших начало советской национальной политике, явившейся становым хребтом всей внутренней и внешней политики советской власти.

В этой связи необходимо сказать также о принятой III Всероссийским съездом Советов 11 и 18 января 1918 г. «Декларации прав трудящегося

---

[1] Цит. из «Истории Советской Конституции в директивах и постановлениях советского правительства 1917—1936 гг.», изд. Ин-та права АН СССР, 1936, стр. 31.

и эксплоатируемого народа» и постановлении того же Съезда «О федеральных учреждениях Российской Республики». В этих исторических документах, написанных Лениным и Сталиным, были изложены принципы организации всей советской государственной системы и установлены «коренные начала федерации Советских Республик» («Декларация»).

В «Декларации прав трудящегося и эксплоатируемого народа» и в постановлении III Съезда «О федеральных учреждениях Российской Республики» от 28 января 1918 г. говорилось, что Российская социалистическая республика учреждается на основе добровольного союза народов России, как федерация советских республик этих народов.

Громадное историческое значение постановления «О федеральных учреждениях» и «Декларации», принятых III Съездом, как и самого III Съезда, обрисовал в своем заключительном слове перед закрытием Съезда В. И. Ленин, подчеркнув то обстоятельство, что «третий съезд Советов открыл новую эпоху всемирной истории», что он закрепил «организацию новой государственной власти, созданной Октябрьской революцией», и что «теперь всем, даже, уверен, нашим врагам видно, что новый строй, власть Советов, не выдумка, не партийный прием, а результат развития самой жизни...».[1]

С полным правом и основанием товарищ Сталин говорил еще в 1922 г., что нам, в стране Советов, «в лагере социализма, удалось с корнем вырвать национальный гнет и наладить взаимное доверие и братское сотрудничество народов».[2]

Товарищ Сталин указывает, что революция в России не победила бы, Колчак с Деникиным не были бы разбиты, если бы русский народ не встретил поддержки со стороны угнетенных народов бывшей Российской империи.

«Но для того, — говорит Сталин, — чтобы завоевать сочувствие и поддержку этих народов, он должен был прежде всего разбить цепи русского империализма и освободить эти народы от национального гнета».[3]

---

[1] Ленин. Соч., т. XXII, стр. 223.
[2] Сталин. Марксизм и национально-колониальный вопрос, 1939, стр. 125.
[3] Сталин. Вопросы ленинизма, стр. 51.

Дружба народов, основанная на подлинном правовом равенстве всех народов Советского Союза, взаимном уважении и доверии, является прочной основой Советского государства.

Товарищ Сталин говорил в докладе «О проекте Конституции Союза ССР», что этот проект исходит из равенства наций и рас, из того, что «разница в цвете кожи или в языке, культурном уровне или уровне государственного развития, равно как другая какая-либо разница между нациями и расами — не может служить основанием для того, чтобы оправдать национальное неравноправие».[1]

Сталинская Конституция, как и первая Союзная Конституция (1924 г.), исходит из того, что все нации и расы, независимо от их прошлого и настоящего положения, от их силы или слабости, должны пользоваться одинаковыми правами во всех областях хозяйственной, общественной, государственной и культурной жизни общества.

Являясь союзным государством, образованным на основе добровольного объединения равноправных советских социалистических республик, СССР всей своей мощью охраняет суверенные права союзных республик и правовое равенство населяющих их народов.

Добровольность и правовое равенство народов Сталин называет основами Советского союзного государства.

Идея сотрудничества и братского сожительства народов нашла свое живое воплощение в Советском Союзе — в Советском многонациональном государстве, представляющем собой братскую семью шестнадцати социалистических республик.

В этой связи уместно остановить внимание на особенностях Советской федерации.

На III Всероссийском съезде Советов (январь 1918 г.) Ленин говорил, что «теперь окончательно признан (в России. — *А. В.*) новый государственный строй Социалистической советской республики, как федерации свободных республик разных наций, населяющих Россию».[2] Ленин говорил в этой речи об особенностях советской федерации, указывая, что основа советской федерации в том, что «мы властвуем, не разделяя, по

[1] Сталин. Вопросы ленинизма, стр. 516.
[2] Ленин. Соч., т. XXII, стр. 223.

жестокому закону древнего Рима, а соединяя всех трудящихся неразрывными цепями живых интересов, классового сознания».[1]

В этом Ленин справедливо видел самую основу советского государственного строя, предсказывая, что вокруг революционной России все больше и больше будут группироваться отдельные различные федерации свободных наций и что «совершенно добровольно, без лжи и железа, будет расти эта федерация, и она несокрушима. Лучший залог ее несокрушимости — те законы, тот государственный строй, который мы творим у себя».[2]

Ленин гениальным своим взором видел рождающийся новый мир социалистической федерации, растущей «на расчищенном от исторического хлама пути». Мы помним эти вдохновенные слова Ленина: «Теперь мы, на расчищенном от исторического хлама пути, будем строить мощное, светлое здание социалистического общества, создается новый, невиданный в истории, тип государственной власти, волей революции призванной очистить землю от всякой эксплуатации, насилия и рабства».[3]

Этот великий пафос вождя пролетарской революции возвышался над всеми трудностями того времени, выражал твердую уверенность поднявшегося на борьбу за свое социальное и политическое освобождение народа в конечном успехе своего дела. Это был пафос новой эпохи, ознаменованной тем, что, как говорил Ленин, «мы действовали без дипломатов, без старых способов, применяемых империалистами», добившись того, что оказался «величайший результат налицо — победа революции и соединение с нами победивших в одну могучую революционную федерацию».[4]

Все развитие Советского государства — Союза Советских Социалистических Республик — доказало величайшую правоту и силу гениального предвидения Ленина.

Народы Советского Союза видят в советской власти родную власть. Соединенные общими интересами, одухотворенные общей идеей, народы Советского Союза с новой силой доказали в годы Отечественной войны свою преданность и любовь к родине-матери.

---

[1] Ленин. Соч., т. XXII, стр. 224.

[2] Там же.

[3] Там же, стр. 223.

[4] Там же, стр. 224.

В одной из своих статей по национальному вопросу, в статье «Политика Советской власти по национальному вопросу в России» (1920 г.), И. В. Сталин говорил: «Советская власть не есть власть, оторванная от народа, — наоборот, она единственная в своем роде власть, вышедшая из русских народных масс и родная, близкая для них».[1] Товарищ Сталин указывал, что «этим, собственно, и объясняется та невиданная сила и упругость, которую обычно проявляет советская власть в критические минуты».[2]

В последующем развитии Советского государства полностью подтвердилось это замечательное предвидение Ленина и Сталина. Советская федерация неизменно росла и развивалась, идя именно по пути, указанному Лениным и Сталиным, все крепче и крепче утверждая государственную связь и государственное единство многочисленных народов бывшей Российской империи, с надеждой и братской любовью поддерживавших героическую борьбу великого русского народа. Многочисленные народы и национальности, населяющие бывшую Российскую империю, справедливо видели в успехе этой борьбы залог своего собственного освобождения, своей собственной государственной самостоятельности и независимости.

Советская автономия — одно из условий прочного государственного союза отдельных советских республик, ибо без этого невозможна «действительная советизация этих областей, превращение их, — как писал товарищ Сталин, — в советские страны, тесно связанные с центральной Россией в одно государственное целое...».[3]

Развивая эту идею, И. В. Сталин указал путь превращения советских стран в «одно государственное целое». Товарищ Сталин показал, что это превращение возможно лишь в виде образования Федерации советских республик, основанной на общности военного и хозяйственного дела. Именно такая федерация, говорил товарищ Сталин, является общей формой государственного союза, которая «может быть прочной, а результаты федерации — действительными лишь в том случае, если она опирается на взаимное доверие и добровольное согласие входящих в нее стран».

[1] Сталин. Марксизм и национально-колониальный вопрос, стр. 82.
[2] Там же, стр. 82—83.
[3] Там же, стр. 83.

Этот важнейший принцип социалистической государственной организации, сформулированный Сталиным в тезисах к X Съезду партии (1921), явился руководящим началом всей советской ленинско-сталинской национальной политики. Этот тезис был развит товарищем Сталиным в докладе на X Всероссийском съезде Советов в 1922 г., посвященном объединению Советских Республик.

В своем докладе И. В. Сталин указал три группы обстоятельств, которые толкают республики на путь объединения, определяют необходимость объединения.

Эти обстоятельства следующие.

Первая группа — внутреннее хозяйственное состояние. Товарищ Сталин относит сюда:

а) Скудость хозяйственных ресурсов, доставшихся нашим молодым республикам в результате семилетней войны (1914—1921 гг.). Эти обстоятельства требовали объединения скудных хозяйственных средств для более рационального их использования и развития основных отраслей хозяйства, составляющих, как сказал товарищ Сталин, «становой хребет советской власти во всех республиках»[1] (подчеркнуто мною. — *А. В.*).

б) Исторически сложившееся естественное разделение труда, хозяйственное разделение труда между различными районами и республиками нашей федерации. Об этом разделении труда И. В. Сталин говорил, что оно делает «невозможным полное развитие отдельных районов при раздельном существовании республик, обязывает республики сплотиться в единое хозяйственное целое».

в) Единство основных средств сообщения по всей федерации; эти средства охарактеризовал товарищ Сталин как «нерв и фундамент всякого возможного объединения».[2]

г) Скудость финансовых средств, вызывающая необходимость сложения финансовых сил отдельных республик в одно целое.

Вторая группа: а) наше военное положение; б) наши внешнеторговые отношения;[3] в) наше дипломатическое положение.

---

[1] См. Сталин. Марксизм и национально-колониальный вопрос, стр. 119.

[2] Там же.

[3] «...наши отношения с заграничным капиталом через Внешторг», как определил эти отношения товарищ Сталин (там же, стр. 120).

Все эти обстоятельства касаются внешнеполитического положения Советских Республик, подвергавшихся в то время организованному бойкоту со стороны капиталистического окружения, стоявших непрерывно перед лицом опасности военного характера, а также и опасности экономической и политической изоляции.

Отсюда задача создать единый неразрывный военный фронт, единый экономический фронт Советских Республик, объединенный фронт по линии дипломатической.

Говоря об этих двух группах обстоятельств, И. В. Сталин подчеркивал особенное значение и исключительную силу этих обстоятельств в условиях перехода после окончания гражданской войны к мирному строительству, когда Республики увидели скудость своих хозяйственных средств и поняли всю необходимость объединения как по внутренней хозяйственной линии, так и по линии внешней.

Товарищ Сталин по этому поводу говорил: «Вот почему теперь, на шестом году существования советской власти, стал на очередь вопрос об объединении независимых советских социалистических республик». [1]

Третья группа фактов, требующих объединения, связана с особенностями советской власти, самая природа которой ведет «к постепенному, но прочному сближению ранее независимых национальностей в одно независимое государство».[2]

Достаточно напомнить об этих трех группах обстоятельств, каждое из которых имеет громадное значение и сыграло выдающуюся роль в решении вопроса об образовании Советской федерации, чтобы не оставалось никакого сомнения в необходимости и целесообразности такой федерации. Заслуга Ленина и Сталина в том, что они указали советским республикам этот путь своего государственного развития, путь, по которому наши республики пошли и который привел их к расцвету и могуществу, выдержавшим испытание огнем и железом в Великую Отечественную войну.

Тогда же были указаны И. В. Сталиным и основы объединения:

а) объединение в федерацию должно быть осуществлено советскими социалистическими республиками — РСФСР, Закавказской Республикой, Украиной и Белоруссией;

---

[1] Сталин. Марксизм и национально-колониальный вопрос, стр. 121.

[2] Там же, стр. 122.

б) объединение должно предусматривать следующие общие народные комиссариаты: внешней торговли, военно-морской, иностранных дел, путей сообщения, связи (почта и телеграф); наркоматы финансов, хозяйства, продовольствия, труда и инспекции должны остаться республиканскими, но действовать по директивам соответствующих наркоматов в союзном центре.

Это — наркоматы, которые впоследствии приобрели название общесоюзных. Такие же наркоматы, как юстиции, просвещения, земледелия, внутренних дел и др. (всего шесть), предлагалось оставить в ведении республиканских правительств. Такое предложение вызывалось необходимостью исходить из особенностей быта, языка, культуры соответствующих договаривающихся республик. Именно эти бытовые и культурные особенности различных республик требовали таких форм государственного управления, которые обеспечивали бы свободу национального развития народов, входящих в состав советских республик.

X Всероссийский съезд Советов признал своевременным объединение Российской Социалистической Федеративной Советской Республики, Украинской Социалистической Советской Республики, Закавказской Социалистической Федеративной Советской Республики и Белорусской Социалистической Советской Республики в Союз Советских Социалистических Республик, положив в основу объединения принцип добровольности и равноправия республик с сохранением за каждой из них права свободного выхода из Союза Республик. Съезд поручил также делегации выработать условия вхождения РСФСР в Союз Республик на началах, соответствующих тем, какие изложены были в тезисах и докладе Сталина, причем подчеркивалась необходимость полного обеспечения интересов национального развития народов договаривающихся республик.

Спустя четыре дня после принятия X Всероссийским съездом этого исторического решения собрался I Съезд Советов Союза ССР. Этот Съезд принял постановление об образовании Союза Советских Социалистических Республик.

Таковы великие принципы ленинско-сталинской национальной политики, на основе которых советская власть сняла так называемый национальный вопрос, не только дав теоретическое решение этой труднейшей проблемы, но и решив его практически, в плане государственного строительства. Гений Ленина и Сталина открыл угнетенным и эксплоатируемым

до советской власти национальностям бывшей Российской империи новый жизненный путь — путь развития национальной культуры, путь развития и творческого роста всех национальных сил многомиллионных народов великой и могучей социалистической державы — Союза Советских Социалистических Республик.

Великими принципами ленинско-сталинской национальной политики Советское государство руководствуется в соответствующих случаях и в области внешнеполитических отношений. Именно Советский Союз последовательно и решительно защищает на международных конференциях права малых и средних народов, их суверенность и независимость, против всех посягательств со стороны империалистических кругов стран буржуазной демократии.

В январе 1946 г. на Лондонской сессии Совета Безопасности советские представители выступили в защиту указанных выше принципов, требуя вывода английских войск из Греции, требуя прекращения войны против индонезийского народа со стороны английских и голландских войск, энергично поддерживая требования Сирии и Ливии о выводе английских и французских войск. Такую же решительную и последовательную позицию Советский Союз занял на Генеральной Ассамблее Организации Объединенных Наций в Нью-Йорке в 1946 г., защищая интересы индусского народа в Южно-Африканском Союзе, в 1947 г.— по палестинскому вопросу и вновь, в 1947 г., — по индонезийскому вопросу, по египетскому вопросу.

Советский Союз последовательно отстаивает в международных отношениях принцип самоопределения народов, добившись того, что этот принцип нашел свое отражение, хотя и в косвенной форме, в уставе Организации Объединенных Наций, отстаивая этот принцип и в практической области внешней политики.

Советский Союз отстаивает независимость и суверенитет народов против всяких попыток ослабить и умалить его значение.

Советский Союз исходит при этом из необходимости «поддержки, решительной и активной поддержки со стороны пролетариата национально-освободительного движения угнетенных и зависимых народов».[1]

---

[1] Сталин. Вопросы ленинизма, стр. 47.

Нельзя не напомнить в качестве наиболее яркой иллюстрации к этому тезису о позиции Советского Союза на Генеральной Ассамблее в 1947 г. в Нью-Йорке по вопросу об учреждении Совета по опеке. Советский Союз разоблачил фальсификаторскую игру, затеянную в этом вопросе англо-американским блоком, пытавшимся в нарушение правил, установленных уставом Организации Объединенных Наций, сохранить за собой господствующее положение на бывших мандатных территориях, превратить их в свои стратегические базы и военные опорные пункты.

Последовательная и глубоко принципиальная позиция Советского Союза в колониально-национальном вопросе привлекает к СССР симпатии угнетенных и зависимых народов и все более и более укрепляет международный авторитет Советского государства, советской внешней политики.

IV

СОВЕТСКОЕ ГОСУДАРСТВО — ГОСУДАРСТВО НОВОГО ТИПА

В предшествующих главах мы уже касались вопросов, непосредственно связанных с учением Ленина — Сталина о государстве и, в частности, о Советском государстве.

Ленин, как это было уже отмечено выше, внес громадный вклад в марксистскую науку о государстве, очистив учение о государстве от фальсификаций оппортунистов, и, что не менее важно, блестяще применил марксистскую теорию в практике Октябрьской революции. Ленин вместе с тем показал, что самое применение этой теории на практике революционного строительства обогащает теорию и двигает ее вперед. Так, в вопросе о сломе государственной машины Ленин показал, что в процессе ликвидации этой машины должны быть сохранены с пользой для пролетариата отдельные ее части, без чего было бы излишне затруднена организация нового государственного механизма.

В этом отношении представляют громадный интерес ленинские идеи, изложенные в таких работах, как, например, статья «Удержат ли большевики государственную власть?».

В этой статье Ленин указывает на то, что в капиталистической системе, в капиталистическом государстве, кроме «преимущественно

«угнетательского» аппарата постоянной армии, полиции, чиновничества», есть аппарат, выполняющий массу работы учетно-регистрационной.

Этого аппарата, говорит Ленин, разбивать нельзя и не надо.

Ленин учит тому, чтобы этот аппарат вырвать из подчинения капиталистам: «от него надо *отрезать, отсечь, отрубить* капиталистов с их нитями влияния, его надо *подчинить* пролетарским Советам, его надо сделать более широким, более всеобъемлющим, более всенародным».[1]

«*Без крупных банков*, — говорит Ленин, — *социализм был бы неосуществим*».[2]

Ленин называет банки «превосходным аппаратом», который «нам *нужен* для осуществления социализма». Капитализм уродует этот превосходный аппарат, от которого нужно отсечь то, что его уродует.

«Единый крупнейший из крупнейших государственный банк, с отделениями в каждой волости, при каждой фабрике — это, — говорит Ленин, — уже девять десятых *социалистического* аппарата. Это — общегосударственное *счетоводство*, общегосударственный *учет* производства и распределения продуктов, это, так сказать, нечто вроде *скелета* социалистического общества».[3]

С этими идеями использования крупнейших банков Ленин связывает идею принудительного синдицирования, т. е. принудительного объединения в союзы под контролем государства. Это, утверждает Ленин, «вполне будет осуществимо в России для Советов, для диктатуры пролетариата, вот что *даст нам «государственный аппарат»* и универсальный, и новейший, и небюрократический».[4]

Мы видим, что «слом» буржуазной государственной машины — это вовсе не акт простого разрушения, как изображают это буржуазные филистеры, это — сложная и продуманная работа, в которой уничтожение «угнетательских» частей сочетается с отбором и использованием

---

[1] Ленин. Соч., т. XXI, стр. 260.

[2] Там же.

[3] Там же.

[4] Там же, стр. 261 — 262. См. также статью «Грозящая катастрофа и как с ней бороться» (стр. 174 и др.).

тех жизненных элементов машины, которые будут включены в новую систему, должны работать, нести свою службу на пользу новому классу, новому обществу.

Это — пример, частность, хотя и имеющая громадное принципиальное значение, бросающая свет на те особенности ленинского учения о государстве, которые характеризуют Советское государство как государство нового типа.

На VIIСъезде партии Ленин дал развернутое изложение особенностей советской государственной системы с точки зрения новизны, с точки зрения того, что Ленин назвал советским типом государства, государством типа Парижской Коммуны. Ленин в черновом наброске проекта программы партии показал эти особенности — это «Десять тезисов о Советской власти», в которых говорится о тех заданиях, которые «исторически на эту форму государственной власти, на этот новый тип государства, ложатся...»[1]

Главное в этих тезисах следующее:

1) объединение и организация рабочих и беднейших крестьян при автоматическом исключении эксплоататорских классов и богатых представителей мелкой буржуазии;

2) их авангард должен воспитывать поголовно все трудящееся население к практическому участию в управлении государством;

3) уничтожение парламентаризма; соединение законодательной и исполнительной государственной работы; слияние управления с законодательством;

4) более тесная связь с массами всего аппарата государственной власти и государственного управления, чем прежние формы демократизма;

5) создание вооруженной силы рабочих и крестьян, наименее оторванной от народа, — один из первых шагов к полному осуществлению вооружения всего народа;

6) более полный демократизм;

7) тесная связь с профессиями и с производительными, экономическими единицами;

8) возможность обойтись без бюрократии, начало реализации этой возможности;

---

[1] Ленин. Соч., т. XXII, стр. 371.

9) не формальное, а практическое осуществление пользования свободой (демократией) для трудящейся и эксплоатируемой массы населения;

10) развитие советской государственности должно итти так, чтобы каждый член Совета обязательно нес постоянную работу по управлению государством, а затем так, чтобы все население поголовно привлекалось постепенно как к участию в советских организациях при условии подчинения организациям трудящихся, так и к несению службы государственного управления.

Таковы задачи Советского государства, поставленные Лениным в годы становления Советского государства.

Что нужно для выполнения этих задач? Ленин ответил на этот вопрос замечательным комментарием, показав *преимущество* (курсив Ленина) Советов в политической, экономической, финансовой и педагогической областях.

Отметим главное: а) распространение советской конституции, *по мере* прекращения сопротивления эксплуататоров, на *все* население.

Эту задачу решила Великая Сталинская Конституция, установившая всеобщее, равное, прямое избирательное право при тайном голосовании;

б) федерация наций, как переход к *сознательному* и более тесному единству трудящихся, научившихся *добровольно* подниматься выше национальной розни. Эту задачу решило образование в 1923 г. Союза Советских Социалистических Республик;

в) обязательно беспощадное подавление сопротивления эксплуататоров; нормы «общей» (*т.-е.* буржуазной) демократии подчиняются этой цели, уступают ей: «свободы» и демократия *не* для всех, а *для* трудящихся;

г) фактическое обеспечение *пользования* свободами со стороны трудящихся;

д) переход *через* Советское государство к постепенному уничтожению государства путем систематического привлечения все большего числа граждан, а затем и *поголовно* всех к управлению государством.[1]

Столь же конкретно Ленин дал указания и в отношении требований в экономической и других областях. Здесь особенно важно отметить требования, касающиеся социалистической организации производства в

[1] См. Ленин. Соч., т. XXII, стр. 372—373.

общегосударственном масштабе — «под общим руководством советской власти, единственно суверенной», а также такие требования, как полное сосредоточение банкового дела в руках государства и всего денежно-торгового оборота в банках, универсализация учета и контроля, организация соревнования «для неуклонного повышения организованности, дисциплины, производительности труда, для перехода к высшей технике, для экономии труда и продуктов, для постепенного сокращения рабочего дня до 6 часов в сутки, для постепенного выравнивания *всех* заработных плат и жалований во *всех* профессиях и категориях».[1]

Такова программа Советского государства как государства социалистического, как государства нового типа. Ленин учил, что выполнение такой программы — дело трудное и что обстановка, в которой необходимо эту программу выполнять, «исключительно сложная», и что пролетариат в выполнении этой программы встретит «напор враждебных сил», перед которым панически готовились отступить всякие новожизненцы,[2] запуганные корниловцами и пугающие рабочих и крестьян.

«Не пугайте, господа, — отвечал им Ленин, — не запугаете...».[3]

Ленин верил и учил всех нас быть твердо уверенными в творческой силе и непобедимости рабочих и крестьян.

«Бояться, — восклицал в негодовании Ленин, — что власть большевиков, то-есть власть пролетариата, которому обеспечена беззаветная поддержка беднейшего крестьянства, «сметут» господа капиталисты! Какая близорукость, какая позорная боязнь народа, какое лицемерие!..»

«Идеи, — добавлял Ленин, — становятся силой, когда они овладевают массами. И именно теперь большевики, т.-е. представители революционно-пролетарского интернационализма, своей политикой воплотили ту идею, которая двигает во всем мире необъятными трудящимися массами».[4]

Советское государство явилось могущественным средством воплощения в СССР идеи социализма. Советское государство явилось «новым типом демократии» (Ленин) в том смысле, о котором говорил Ленин в своих десяти тезисах и в целом ряде других важнейших документов и трудов.

[1] Ленин. Соч., т. XXII, стр. 374.
[2] Из газеты «Новая жизнь».
[3] Ленин. Соч., т. XXI, стр. 272.
[4] Там же, стр. 277, 279.

Одной из важных черт «новой демократии» явилось не только отрицание формального равенства буржуазии и пролетариата, бедных и богатых, эксплоататоров и эксплоатируемых и выключение буржуазии из «демократии» и «свобод», примером чего является лишение эксплоататорских элементов избирательного права, и не только подавление этих элементов. Важной чертой советской государственности, именно как государственности нового типа, является самый метод подавления. Это в высшей степени важное обстоятельство, создающее принципиально качественное различие между этими основными функциями социалистического государства и государства любой другой общественной формации. Ленин учит, что подавлять эксплоататоров нельзя полицией — «их может подавлять только сама масса, аппарат должен быть связан с массами, должен ее представлять, как советы».[1]

Ленин учит, что «было бы величайшей глупостью и самым вздорным утопизмом полагать, что без принуждения и без диктатуры возможен переход от капитализма к социализму».[2]

Ленин и Сталин учат, что это принуждение Советское государство применяет при участии народных масс, усилиями самих масс. По Ленину, диктатура пролетариата направляет свои удары против эксплоататоров через массы трудящихся, через организации, «построенные так, чтобы именно такие массы будить, поднимать к историческому творчеству (советские организации принадлежат к этого рода организациям)».[3]

Сталин подчеркивает, что репрессии, — этот обязательный элемент принуждения, — являются необходимым элементом наступления, но элементом вспомогательным, а не главным.

Сталин в борьбе, например, с расточительностью наших управляющих органов и с расточительностью в быту призывал в первую очередь к созданию «атмосферы общего морального бойкота и ненависти окружающей публики» к расточителям, такой атмосферы, которая «делала бы невозможным и жизнь и существование воров и расхитителей народного добра».[4]

Нравственное воздействие, перевоспитание, исправление, прививание

---

[1] Ленин. Соч., т. XXII, стр. 353.

[2] Там же, стр. 458.

[3] Там же, стр. 462.

[4] «Правда», № 89 от 18 апреля 1926 г.

4*

общественных навыков и традиций — вот элементы того принуждения в широком смысле этого слова, которое в руках советских учреждений и органов государственной власти становится могучим средством регулирования общественных отношений, борьбы с правонарушениями и правонарушителями и даже преступниками.

Ленин требовал поэтому поголовного привлечения трудящихся к отправлению правосудия, подчеркивая, что на советские суды ложится громадная задача воспитания населения в духе трудовой дисциплины, что «граждане должны участвовать поголовно в суде и в управлении страны».[1]

Советское государство — государство нового типа, ибо оно дает высшую форму организации демократии, дает образец народного государства, построенного и приводящегося в движение руками самого народа, т. е. рабочих, крестьян, интеллигенции, под руководством самой передовой партии — партии Ленина — Сталина.

Отвечая на записки на Чрезвычайном Всероссийском железнодорожном съезде в марте 1918 г., Ленин говорил: «... Советская власть дает не таких представителей, которые фехтуют в парламентах и обмениваются блестящими речами, создавая прочное господство капитала и чиновничьего аппарата. А Советская власть идет от самих трудящихся масс, она дает не парламент, а собрание трудовых представителей, которое издает законы, исполняющиеся непосредственно, проходящие в жизнь и ставящие своей задачей борьбу с эксплуататорами».[2] «Нам нужно представительство не буржуазное, а представительство эксплуатируемых и угнетенных, которое будет вести беспощадную борьбу с эксплуататорами. Вот какое намерение Советской власти; в него не входят ни парламент, ни референдум».[3]

Тридцать лет Советского государства завершают величественную эпоху в истории советской власти — эпоху построения социализма и утверждения подлинной, самой полной и последовательной в мире демократии в гигантской стране, в первоклассной и могущественной социалистической державе.

---

[1] Ленин. Соч., т. XXII, стр. 354.
[2] Там же, стр. 239.
[3] Там же, стр. 240.

Тридцать лет Советского государства — это славный героический путь рабочих и крестьян нашей страны, ведомых партией Ленина — Сталина к коммунизму, в мир полного и окончательного торжества идеи равенства и братства народов.

V

УЧЕНИЕ СТАЛИНА О СОЦИАЛИСТИЧЕСКОМ ГОСУДАРСТВЕ

Победа социализма в СССР в результате ликвидации эксплоататорских классов и прочного установления государственной и общественной собственности, основа всего политического и общественного советского строя, выдвинула ряд проблем в области учения о государстве. Эти теоретические проблемы имели к тому же серьезное практическое значение, требовали ответа. Важнейшая из этих проблем касалась судьбы самого государства, касалась вопроса о том, оправдано ли при отсутствии в стране враждебных классов, в условиях движения нашего общества к коммунизму, существование государства. Правильно ли, что в указанных условиях мы не содействуем отмиранию социалистического государства, а, наоборот, видим задачу дальнейшего укрепления нашего государства?

Таким образом, в порядке дня стали вопросы о судьбах социалистического государства и о так называемом отмирании государства.

Самая постановка этих вопросов свидетельствовала о недоработанности и недостаточности некоторых общих положений учения о государстве.

Так объяснил товарищ Сталин возникновение новых проблем в учении о государстве, присоединив к сказанному в качестве объяснения также вполне обоснованное соображение о том, что авторы этих вопросов не разобрались в исторических условиях, когда вырабатывались отдельные положения марксистского учения о государстве, какова международная обстановка, в которой находится Советское государство, не поняли и недооценили факта капиталистического окружения, со всеми вытекающими из этого факта последствиями, вроде засылки в СССР банд шпионов, разведчиков, убийц и вредителей.

Вопросы социалистического государства, в связи с указанными фактами, встали во весь рост, потребовав серьезного к себе внимания, так как они, далеко выходя из рамок только теоретического интереса, приобретали крупнейшее практическое значение.

На эти вопросы ответил в своем выступлении на XVIII Съезде партии товарищ Сталин, дав развернутое учение о социалистическом государстве и тем самым двинув вперед марксистско-ленинскую теорию в этой в высокой степени важной, решающей области научного коммунизма.

Великая заслуга Сталина в том, что он продолжил и завершил начатую Лениным, но, вследствие преждевременной смерти нашего великого учителя, не оконченную им работу по теории государства.

«Не может быть сомнения, — говорил Сталин в своем докладе на съезде, — что Ленин имел в виду во второй части своей книги («Государство и революция». — *А. В.*) разработать и развить дальше теорию государства, опираясь на опыт существования Советской власти в нашей стране. Но смерть помешала ему выполнить эту задачу. Но чего не успел сделать Ленин, должны сделать его ученики», сказал Сталин. Сталин это сделал, открыв новую страницу в теории государства, и социалистического государства в особенности.

Сталин показал, что «две основные функции характеризуют деятельность государства: внутренняя (главная) — держать эксплоатируемое большинство в узде и внешняя (не главная) — расширять территорию своего, господствующего класса за счет территории других государств, или защищать территорию своего государства от нападений со стороны других государств».[1]

Эта функция была у рабовладельческого государства, как и у капиталистического. Стремление капиталистического государства к расширению своей территории, к экспансии вытекает из всего характера, из самой системы капитализма. Это настолько очевидно, что даже Притт в своем докладе в Брюсселе 3.VI с. г. должен был это признать в отношении, в частности, США, когда сказал, что при всей своей силе США очень слабы и им, как всем государствам, достигшим высшей стадии капитализма, необходима экспансия. Притт с полным основанием добавил к этому, что «руководящие круги американского империализма опьянены

[1] Сталин. Вопросы ленинизма, стр. 604.

своей мощью и особо опасны потому, что они не обладают техникой управления миром».

Сталин учит, что Советское государство, несмотря на то, что оно стало в России на место разрушенной государственной машины буржуазии, сохранило некоторые функции старого государства, измененные применительно к потребностям пролетарского государства. Важно отметить, что некоторые функции старого государства пролетарское государство сохраняет в **измененном** виде, причем эти изменения **определяются потребностями** пролетарского государства. Известно, вместе с тем, что потребности пролетарского государства в свою очередь определяются его, так сказать, природой, т. е. теми его качествами, свойствами, особенностями, принципами, которые присущи пролетарскому государству.

Сталин учит, что формы и первоначальные функции Советского государства не могут остаться неизменными.

«На самом деле формы нашего государства меняются и будут меняться в зависимости от развития нашей страны и изменения внешней обстановки».[1]

Сталин приводит при этом указание Ленина о том, что в течение всего «*исторического периода*, отделяющего капитализм от «общества без классов», от коммунизма», неизбежно громадное обилие и разнообразие политических форм, но «сущность будет при этом неизбежно одна: *диктатура пролетариата*».[2]

Сталин учит, что наше социалистическое государство со времени Октябрьской революции прошло в своем развитии две главные фазы. Учение Сталина о двух фазах развития социалистического государства имеет особенно важное значение, так как оно не только содержит объяснения, какие именно функции и в силу каких именно исторических условий и причин осуществляло в первую и во вторую фазы Советское государство. Это учение о фазах важно и потому, что оно объясняет неизбежность дальнейшего развития социалистического государства, неизбежность новых изменений его функций.

«Первая фаза — это период от Октябрьской революции до ликвидации эксплоататорских классов. Основная задача этого периода состояла

---

[1] Сталин. Вопросы ленинизма, стр. 604—605.
[2] Ленин, Соч., т. XXI, стр. 393.

в подавлении сопротивления свергнутых классов, в организации обороны страны от нападения интервентов»[1] и пр.

Сталин учит, что во вторую фазу развития Советского государства «отмерла функция военного подавления внутри страны, ибо эксплоатация уничтожена, эксплоататоров нет больше и подавлять некого».[2]

Но, говорит Сталин, вместо этой функции появилась функция охраны социалистической собственности от воров и расхитителей народного добра, сохранилась полностью функция военной защиты страны от нападений извне, что означает сохранение Советской Армии, Военно-Морского Флота, карательных органов, разведки, необходимых «для вылавливания и наказания шпионов, убийц, вредителей, засылаемых в нашу страну иностранной разведкой».

«Сохранилась и получила полное развитие функция хозяйственно-организаторской и культурно-воспитательной работы государственных органов».[3]

Сталин особо подчеркивает в своем учении о социалистическом государстве, что острие нашей армии, карательных органов и разведки теперь обращено не внутрь страны, а вне ее, против внешних врагов.

Таково социалистическое государство, таковы его функции во второй фазе развития.

Сталинское учение о социалистическом государстве, изложенное нашим вождем на XVIII Съезде партии (1939), является развитием идей, выраженных в труде Сталина «О диалектическом и историческом материализме» (1938). Противопоставляя социалистический строй в СССР капиталистическому строю, Сталин подчеркивал, что в СССР взаимные отношения людей в процессе производства характеризуются как отношения товарищеского производства и социалистической взаимопомощи свободных от эксплоатации работников. Должно быть ясно, что характер и особенности социалистического производства неизбежно определяют характер и особенности всех общественных отношений, а следовательно также определяют характер и особенности государственно-правовых отношений самого государства и его функций.

Значение сталинского учения о социалистическом государстве яв-

[1] Сталин. Вопросы ленинизма, стр. 605.
[2] Там же, стр. 606.
[3] Там же.

ляется тем более исключительным, что оно кладет конец недооценке роли и значения нашего государства, его механизма, его разведки, которые-де скоро придется сдать в музей древностей. В этой болтовне о том, что Советское государство придется сдать в музей древностей, слышатся отголоски вредительско-провокаторских «теориек» троцкистско-бухаринских изменников. Известно, что эти господа не мало старались, по заданиям иностранных разведок, расшатать Советское государство, внушая тлетворные мысли о «растворении», «засыпании», «погружении в экономике» Советского государства, о том, что Советское государство существует лишь в психике людей, что советский закон и советское право — не что иное, как «юридическое барахло», и тому подобный провокаторский вздор.

В свое время эти «теорийки» преображенских, сокольниковых, бухариных были до конца разоблачены, причем главная заслуга в этом разоблачении принадлежит Сталину, выступившему с такими, сыгравшими историческую роль работами, как речь «О правом уклоне в ВКП(б)» (1929), «Год великого перелома», «К вопросам аграрной политики в СССР» и др.

Сталин тогда же разоблачил антимарксистский и вредный для социалистического строительства характер «упражнений» Бухарина и К°, направленных на то, чтобы выбить из рук пролетариата такой инструмент борьбы с эксплоататорскими элементами, как Советское государство. Не случайно ведь Бухарин доказывал, что кулацкий террор против советских людей прекратится в результате улучшения советского аппарата и что, следовательно, кулаки стреляют в передовых советских людей по той простой причине, что государственный аппарат работает плохо.

Что касается вопроса о государстве, то верхом извращений марксизма-ленинизма явилась пропаганда отмирания Советского государства, для чего была даже придумана специальная формула, согласно которой «сперва отомрут армия и флот», «потом — система карательных и репрессивных органов», далее — «принудительный труд». Эта «формула-схема» была, как известно, сокрушительно разгромлена Лениным, заметившим: «Не наоборот-ли: сначала «далее», затем «потом» и наконец «сперва»?».

Ясно было, что такие «схемы» в условиях острой классовой борьбы переходного периода от капитализма к коммунизму могли изобретать либо сумасшедшие, либо провокаторы и предатели. В 1936—1938 гг. выяснилось, что авторами таких «схем» были, действительно, предатели, какими стали, в конце концов, троцкисты и бухаринцы, превратившиеся

«в оголтелую и беспринципную банду вредителей, диверсантов, шпионов и убийц, действующих по заданиям разведывательных органов иностранных государств».[1]

Вопрос об «отмирании» государства, основательно запутанный Бухариным еще в 1916 г., являлся в руках предателей и двурушников весьма удобным инструментом подрыва мощи Советского государства. Он был особенно удобен тем, что легко служил двурушникам для прикрытия их действительных, подлых замыслов. Ведь проблема «отмирания» государства — вполне теоретическая проблема; эта проблема — самая что ни на есть научная; эта проблема — самая «революционная», самая «левая», наиболее импонирующая чувству ненависти к традиционному государству, исторически сложившемуся как аппарат угнетения и подавления. И надо сказать, что двурушники эту «проблему» ловко использовали в своих предательских целях, пытаясь отравить сознание наших людей ядом отрицания социалистического характера Советского государства, советских банков, советской промышленности и торговли, советского закона, суда, карательных органов.

Эти попытки встретили решительный отпор со стороны нашей партии и не дали ожидавшихся предателями результатов. Рецидива не будет. Но «непозволительно беспечное отношение» к вопросам теории государства, как указывал на XVIII Съезде партии И. В. Сталин, привело к недооценке роли Советского государства и вызвало необходимость разъяснения некоторых вопросов, связанных с проблемой развития социалистического государства. Это с честью выполнил Сталин. Развивая далее марксистско-ленинскую теорию социалистического государства, в частности в важнейшем вопросе об его отмирании, Сталин показал, что формула Энгельса в «Анти-Дюринге» о том, что «государство (на известной стадии своего развития. — *А. В.*) не «отменяется», а оно отмирает», — правильна, но при одном из двух условий. Указание на эти условия, — чего не дал Энгельс, да и не мог дать, по той причине, что, как разъяснил товарищ Сталин, при исследовании этого вопроса Энгельс отвлекался от такого фактора, как международные условия, международная обстановка, — означало новый, громадный шаг вперед в теории социалистического государства.

---

[1] Сталин. О недостатках партийной работы и мерах ликвидации троцкистских и иных двурушников.

Каковы эти условия? Их два, по учению Сталина: первое — если брать социалистическое государство изолированно от международных отношений, с точки зрения лишь внутреннего развития страны; второе — если предположить, что социализм победил во всех странах или в большинстве стран, что налицо, вместо капиталистического окружения, социалистическое окружение и что нет угрозы нападения извне, нет больше нужды в усилении армии и государства.

«Ну, а если, — продолжает Сталин развивать свою теорию, — социализм победил только в одной, отдельно взятой стране, и отвлекаться, ввиду этого, от международных условий никак невозможно, — как быть в таком случае?»[1]

Сталин дал ответ и на этот вопрос, указав на то, что страна, в которой победил социализм, которая находится в капиталистическом окружении и которая подвержена угрозе нападения извне, должна иметь свое государство — «достаточно сильное государство, — для того, чтобы иметь возможность защищать завоевания социализма от нападения извне».[2]

Сталин видит далеко вперед. Он видит неизбежность дальнейшего развития социалистического государства. Он видит необходимость сохранения социалистического государства до тех пор, пока не будет ликвидировано капиталистическое окружение и не будет заменено окружением социалистическим. Таково сталинское учение о социалистическом государстве — о фазах его развития, о его функциях, об условиях его жизни и деятельности на путях к коммунизму, о его задачах и методах борьбы за достижение конечных целей, за построение коммунистического общества. Сталинское учение о социалистическом государстве завершило развитие марксистско-ленинской теории о государстве; оно вооружило строителей социализма ясным пониманием перспективы будущего, дало им в руки могучее орудие дальнейшего успешного строительства и борьбы за бесклассовое общество, за коммунизм.



Тридцатилетие социалистической революции трудящиеся нашей страны встречают на подъеме своих творческих сил.

[1] Сталин. Вопросы ленинизма, стр. 602—603.
[2] Там же, стр. 603.

Великая Отечественная война явилась историческим испытанием всех материальных и духовных сил советского народа и еще более укрепила мощь Советского Союза. Война показала неисчерпаемые резервы сил и средств советской страны, способной не только преодолеть, казалось бы, непреодолимые военные трудности, но, быстро оправляясь от понесенных жертв и причиненного ей врагом громадного ущерба, подняться на следующую ступень своего развития, неуклонно двигаясь вперед и вперед.

Окончание войны, победоносно завершенной вооруженными силами союзников, и переход от войны к миру поставили перед Советским Союзом новые задачи как внутри страны, так и в области внешнеполитических отношений. Первый послевоенный план — план восстановления народного хозяйства, ликвидации разрушений и бедствий, причиненных советской стране войной, навязанной советскому народу разбойничьей гитлеровской Германией, требует гигантских усилий для своего выполнения. Движимые чувством подлинного советского патриотизма и преисполненные твердой решимости в кратчайший срок выполнить и перевыполнить этот план, советские люди ведут сейчас героическую борьбу на фронте хозяйственного и культурного строительства. Они мобилизуют все свои силы и всю энергию для решения этой важнейшей государственной задачи, смело и уверенно глядя в будущее, твердо и мужественно идя вперед под руководством своего Правительства во главе с великим Сталиным.

Советский народ уверен в успехе своего дела, уверен в том, что он быстро залечит раны, нанесенные ему жестокой войной, и с еще большей силой развернет свои творческие способности, обеспечив новый, еще более могучий расцвет Советского государства.

Советский народ ценит прочный демократический мир как основу прогресса и возможность неуклонно улучшать благополучие миллионных народных масс.

Советскому народу и Советскому государству, являющемуся плотью от плоти народа, чужды милитаристские завоевательные стремления. Политика Советского государства — это политика труда и мира, дружбы и сотрудничества между народами.

Советский народ всегда готов к обороне, к сокрушительному отпору преступным проискам поджигателей войны, всегда готов со всей реши-

тельностью ответить на удар беспощадным ударом. Советские люди знают, что их путь впереди усеян многочисленными препятствиями и трудностями, что у них не мало врагов, ожидающих удобного времени и подходящих условий, чтобы попытаться сокрушить страну социализма, могучую цитадель демократии и мира.

Советский Союз, несмотря на многие неблагоприятные обстоятельства, принял предложение и вступил в свое время в Лигу Наций, дававшую некоторую, незначительную, возможность противостоять угрожавшей тогда миру агрессии со стороны фашистских государств, пользовавшихся поддержкой мюнхенцев. Советский Союз, верный своей политике мирного сотрудничества и дружбы между народами, явился одним из инициаторов и главных организаторов новой международной организации — Организации Объединенных Наций, видя в ней новый инструмент международного сотрудничества больших и малых стран в целях обеспечения мира и безопасности народов.

В течение двух лет, протекших со дня учреждения Организации Объединенных Наций, Советский Союз неуклонно и последовательно борется за принципы международного сотрудничества, не щадя усилий для обеспечения успеха в этом деле. Он действует в соответствии с принципами, провозглашенными в уставе этой организации, добиваясь от всех ее участников полного и безусловного соблюдения этих принципов.

Окончание второй мировой войны и переход на мирные рельсы, к сожалению, оказались связанными с обострением ряда серьезных разногласий между союзниками. Приглушенные войной и общей задачей сокрушить врага, эти противоречия вскрылись с новой силой, придав тревожный характер внешнеполитическим отношениям между великими державами. Окончание войны развязало новые агрессивные стремления некоторых стран, преувеличивающих свои силы и роль в международных отношениях. Среди этих стран главное место занимают Соединенные Штаты Америки, направляющие свои усилия к тому, чтобы обеспечить себе привилегированное положение среди остальных государств, опираясь на свою экономическую мощь и пользуясь в соответствующих случаях для осуществления своих планов методами политического и экономического давления. «Доктрина» Трумэна и «план» Маршалла, так же как и попытки практического претворения их в жизнь, сделанные на Парижской конференции в июле 1947 г., являются

наиболее яркими доказательствами того «нового курса» внешней политики, который в Соединенных Штатах Америки сменил старый курс, проводившийся Рузвельтом и его сторонниками в интересах укрепления мирного международного сотрудничества. Новый курс внешней политики США, блокирующихся с Великобританией, Францией и некоторыми другими странами, полностью включившимися в орбиту американского влияния, чреват серьезными опасностями для всеобщего мира. Острие этого курса направлено на обеспечение успеха экспансионистским планам реакционных кругов США и других капиталистических стран. Оно направлено против интересов стран новой демократии Восточной Европы и, прежде всего, против интересов Советского Союза, социалистического государства рабочих и крестьян.

Непоколебимо, как скала, противостоит этому «курсу» Советский Союз, страна социализма, страна подлинного народоправства, подлинной демократии, верный, последовательный и бесстрашный хранитель демократических прав и свобод народов, демократического и прочного мира во всем мире, страж независимости и безопасности миролюбивых наций, поборник равенства прав и взаимного уважения интересов больших, средних и малых государств, готовых к сотрудничеству во имя прогресса всего человечества.

В трудных условиях современного международного положения Советский Союз продолжает непоколебимо вести свою внешнюю политику, опирающуюся на великие принципы суверенного равенства народов, защиты прочного демократического мира и безопасности, содействия экономическому и социальному прогрессу всех государств, на уважение принципов справедливости и международного права, как основы дружественного международного сотрудничества больших и малых стран.

МАТЕМАТИКА



Академик
И. М. Виноградов

АДДИТИВНЫЕ ПРОБЛЕМЫ ТЕОРИИ ПРОСТЫХ ЧИСЕЛ



ри целом положительном r будем рассматривать какую-либо совокупность Ω целых точек $(x_1, \ldots, x_r)$ r-мерного пространства (*целыми* мы называем такие точки, у которых все координаты суть целые числа). Эту совокупность Ω можно составлять самыми разнообразными способами: можно составить ее из всех целых точек r-мерного пространства, но можно составить ее и из любой части этих точек, например: из точек, у которых все координаты положительные; из точек, у которых координаты не делятся одновременно ни на какое целое число, превосходящее единицу; из точек, у которых все координаты суть простые числа, и т. д.

Будем далее рассматривать какую-либо функцию $f(x_1, \ldots, x_r)$, имеющую для точек совокупности Ω целочисленные значения; в частности, в качестве функции $f(x_1, \ldots, x_r)$ можно рассматривать любую целую рациональную функцию, принимающую для целочисленных значений $x_1, \ldots, x_r$ целочисленные значения. Например, можно положить $f(x_1, \ldots, x_r) = x_1^n + \ldots + x_r^n$, где n — целое положительное число.

Соответствующие точкам совокупности Ω значения функции $f(x_1, \ldots, x_r)$, будучи целыми числами, каким-то образом распределятся среди всех целых чисел. Изучение законов такого распределения — одна из важнейших проблем теории чисел. В частности, весьма важно изу-

чение поведения функции $I(N)$, выражающей число точек совокупности Ω, для которых $f(x_1, \ldots, x_r)$ равно заданному целому числу N, иными словами — функции $I(N)$, выражающей число представлений целого числа N в форме

$$N=f(x_1, \ldots, x_r).$$

Изучение функции $I(N)$ при тех или иных ограничениях, налагаемых на область Ω и функцию $f(x_1, \ldots, x_r)$, привело в разное время ученых к постановке ряда важных и глубоких проблем. Мы дадим сейчас несколько примеров таких проблем.

Пусть, например, совокупность Ω состоит из всех целых точек (x, y) плоскости и $f(x, y)=x^2+y^2$. При большом положительном a можно поставить вопрос об асимптотической (приближенной) формуле для суммы

$$T(r)=\sum_{0 \leqslant N \leqslant a^2} I(N)$$

всех значений функции $I(N)$, отвечающих неотрицательным значениям N, не превосходящим a^2. Очевидно, $T(a)$ есть не что иное, как число целых точек (x, y) плоскости, лежащих внутри или на контуре круга радиуса a, центр которого лежит в начале координат. Здесь тривиально можно установить, что $T(a)$ асимптотически (приближенно) равно площади πa^2 круга, причем остаточный член R асимптотической формулы

$$T(a)=\pi a^2+R$$

при неограниченном возрастании a растет как величина порядка не выше a, т. е. $R=O(a)$. Естественно возникает вопрос о нетривиальной и притом возможно более точной оценке для R. Видоизменяя метод Вороного (уже примененный в 1903 г. этим ученым к другому аналогичному вопросу), Серпинский дал первую нетривиальную оценку для R; он показал, что $R=O\left(a^{\frac{2}{3}}\right)$. Эта оценка была далее еще улучшена ван-дер-Корпутом, который показал, что $R=O\left(a^{\frac{2}{3}-\delta}\right)$, где δ — положительное, хотя и очень малое постоянное. После этого значительных продвижений в этом вопросе не было; правда, существует гипотеза, что $R=O\left(a^{\frac{1}{2}+\varepsilon}\right)$, где ε — произвольно малое положительное постоянное.

Рассмотренный вопрос есть частный случай общей проблемы об асимптотической формуле для числа целых точек данной плоской области, уже решенной (мои работы 1917 г. и более поздние, а также не зависимые от моих работ работы ван-дер-Корпута) для весьма широкого класса областей. Аналогичная проблема решена также и для довольно широкого класса областей обычного трехмерного пространства, а также для многих областей, принадлежащих пространствам более чем трех измерений. Однако, подобно рассмотренной выше проблеме об асимптотической формуле для числа целых точек области круга, полученные результаты в отношении точности оценки остаточного члена еще весьма далеки от окончательных.

Пусть совокупность Ω состоит из целых точек (x, y) плоскости с взаимно простыми x и y и пусть опять $f(x, y)=x^2+y^2$. Но теперь мы будем рассматривать самое число $I(N)$, т. е. число представлений целого числа N в форме

$$N=x^2+y^2 \tag{1}$$

с взаимно простыми x и y. Этим вопросом занимался еще Ферма, при этом он высказал утверждение, что число N представимо в форме (1), если оно — простое число вида $4m+1$. Утверждение Ферма было потом доказано Эйлером. Из результата Эйлера следует и такой общий результат: *необходимым и достаточным условием представимости целого числа* $N>2$ *в форме* (1) *является отсутствие среди его делителей числа* 4 *и простых чисел вида* $4m+3$. Более того, при выполнении этого условия можно указать и точную формулу для $I(N)$; она будет: $I(N)=2^{k+3}$, где k — число различных нечетных простых делителей числа N.

Дальнейшим развитием рассмотренного вопроса явилась проблема изучения числа $I(N)$ представлений целого числа N в форме

$$N=ax^2+2\,bxy+cy^2$$

(a, b и c — целые); глубокое исследование этой проблемы дает классическая теория бинарных квадратичных форм. Неизмеримо более трудной является проблема изучения числа $I(N)$ представлений целого числа N в форме

$$N=a_n x^n+\ldots+a_1 xy^{n-1}+a_0 y^n$$

5*

при условии, что $a_n z^n + \ldots + a_1 z + a_0$ — неприводимый полином степени $n > 2$ с целыми $a_n, \ldots, a_1 a_0$ и с a_n не равным нулю. Однако эта проблема еще весьма мало разработана, и крупнейшим достижением в ней является лишь теорема (*Thue*, *Siegel*) о том, что число $I(N)$ не может быть бесконечным.

Пусть совокупность Ω состоит из целых точек $(x_1, \ldots, x_r)$ r-мерного пространства с неотрицательными координатами и пусть $f(x_1, \ldots, x_r) = x_1^n + \ldots + x_r^n$, где n — целое, превосходящее 2. Варинг еще в 1770 г. высказал следующее утверждение: каждому целому n, превосходящему 2, отвечает свое r с условием, что для всякого целого неотрицательного N имеем $I(N) \geqslant 1$, т. е. с условием, что всякое целое неотрицательное N представимо в форме

$$N = x_1^n + \ldots + x_r^n \tag{2}$$

с целыми неотрицательными $x_1, \ldots, x_r$. Это утверждение получило название *проблемы Варинга*; оно было доказано Гильбертом в 1909 г. Однако метод, которым Гильберт достиг цели, был весьма несовершенен. Харди и Литтльвуд в 1919 г. разработали для той же цели другой метод, несравнимо более могущественный. Проблема Варинга принадлежит к числу так называемых *„аддитивных проблем“* теории чисел (т. е. проблем, касающихся представления целых чисел суммами целочисленных слагаемых наперед заданного вида), и в основу своего метода Харди и Литтльвуд положили уже применявшуюся в свое время к простейшим аддитивным проблемам идею Эйлера рассматривать коэффициенты ряда, получаемого перемножением степенных рядов. С помощью этого метода Харди и Литтльвуд не только по-новому доказали утверждение Варинга, но и значительно уточнили вопрос о верхней границе для функции $G(n)$ (эта функция обозначает наименьшее r, для которого существует какое-либо число N_0 с условием, что для всякого $N \geqslant N_0$ имеем $I(N) \geqslant 1$). Именно при $n > 2$ они доказали, что

$$G(n) \leqslant (n-2)\, 2^{n-2} h, \tag{3}$$

где при неограниченном возрастании n число h неограниченно приближается к единице. Более того, при

$$r \geqslant (n-2)\, 2^{n-1} + 5 \tag{4}$$

Харди и Литтльвуд вывели следующую асимптотическую формулу:

$$I(N)=\frac{\Gamma(1+\nu)^r}{\Gamma(r\nu)}N^{r\nu-1}\mathfrak{S}+O(N^{r\nu-1-\delta}), \qquad (5)$$

где $\nu=n^{-1}$, δ — положительное число, зависящее только от n, а $\mathfrak{S}$ — так называемый „особый ряд". При этом Харди и Литтльвуд показали, что $\mathfrak{S}>c$, где c — положительное постоянное, зависящее только от n и r. Впоследствии, путем усовершенствования метода Харди и Литтльвуда было показано (*Ниа* в 1938 г.), что правые части неравенств (3) и (4) могут быть заменены числом 2^n+1. Следует отметить, что незадолго до создания Харди и Литтльвудом их метода Вейль (в 1914 г.) разработал метод оценки суммы вида

$$\sum_{(\Lambda)} e^{2\pi i\, F(x_1,\,\ldots,\,x_k)}\,, \qquad (6)$$

где k — целое положительное число (особенно важен случай $k=1$) и суммирование распространяется на все точки $(x_1,\ldots,x_k)$, принадлежащие совокупности Λ. Предполагается, что совокупность Λ состоит из всех без исключения целых точек некоторой k-мерной области; на эту k-мерную область и на функцию $F(x_1,\ldots,x_k)$ должны быть наложены некоторые ограничения. Оценки сумм (6) играют важную роль в применениях метода Харди и Литтльвуда (так, например, оценки сумм вида

$$\sum_{x=0}^{P-1} e^{2\pi i\alpha x^n}$$

играли важную роль в работах Харди и Литтльвуда, посвящённых проблеме Варинга). Мною (в 1934 г.) был разработан новый метод в аналитической теории чисел, дающий, в частности, для сумм (6) оценки несравнимо более точные, нежели оценки, получаемые по методу Вейля. Применение моего метода позволило вместо верхней границы (3) для $G(n)$ получить такую:

$$G(n)<n(3\log n+11).$$

Эта граница при возрастании n растет как величина порядка $n\log n$, и в этом отношении она близка к окончательной. Действительно, можно легко показать, что $G(n)>n$. Существует гипотеза, что

$$G(n)\leqslant cn,$$

где c — положительное постоянное (не зависящее от n). Применение моего метода позволило также вместо нижней границы (4) для чисел r, при которых справедлива асимптотическая формула (5), получить такую:

$$r \geqslant [10\, n^2 \log n].$$

Но эта последняя граница, вероятно, еще очень далека до окончательной. Существует гипотеза, что асимптотическая формула (5) справедлива при

$$r \geqslant c_1 n.$$

Метод Харди и Литльвуда оказался применимым к весьма широкому классу аддитивных проблем. Мы остановимся сейчас на кратком описании этого замечательного метода и именно того варианта этого метода, который применяется в моих работах.

Пусть задана какая-либо совокупность Ω целых точек $(x_1, \ldots, x_r)$ r-мерного пространства и пусть $f(x_1, \ldots, x_r)$ — какая-либо функция, принимающая в точках совокупности Ω целочисленные значения. Легко показать, что интеграл

$$\int_0^1 e^{2\pi i \alpha H} d\alpha,$$

где H — целое число, равен 1 в случае, когда $H=0$, и равен 0 в противном случае. Отсюда легко следует, что функция $I(N)$ может быть представлена интегралом

$$I(N) = \int_0^1 \sum_{(\Omega)} e^{2\pi i \alpha (f(x_1, \ldots, x_r) - N)} d\alpha, \tag{7}$$

где суммирование $\sum\limits_{(\Omega)}$ распространяется на все точки совокупности Ω.

Далее мы выбираем два числа τ_0 и τ с условиями $1 < \tau_0 < \tau < N$. Из теории непрерывных дробей известно, что всякое α интервала

$$-\frac{1}{\tau} \leqslant \alpha \leqslant 1 - \frac{1}{\tau} \tag{8}$$

(которым, очевидно, можно заменить интервал $0 \leqslant \alpha \leqslant 1$ интегрирования в равенстве (7)) можно представить в форме

$$\alpha = \frac{a}{q} + z, \qquad -\frac{1}{q\tau} \leqslant z \leqslant \frac{1}{q\tau} \tag{9}$$

с условием, что a и q взаимно простые и $0 < q \leqslant \tau$. Из интервала (8) мы выделим *основные интервалы*, включающие все α вида (9) с условием $q \leqslant \tau_0$ (интервалы, отвечающие малым q); при этом часто неравенства, указанные для z, заменяют и такими:

$$-\frac{1}{\tau} \leqslant z \leqslant \frac{1}{\tau}.$$

Число τ_0 предполагается выбранным таким образом, чтобы основные интервалы, отвечающие различным парам чисел a и q, не могли содержать общих значений α. Среди основных интервалов будем иметь интервал, отвечающий случаю $a=0$, $q=1$; из (9) следует, что это будет интервал

$$-\frac{1}{\tau} \leqslant \alpha \leqslant \frac{1}{\tau}.$$

Интервалы, остающиеся после выделения из интервала (8) основных интервалов, мы назовем *дополнительными интервалами* (интервалы, отвечающие большим q). Каждое α, принадлежащее дополнительному интервалу, очевидно, представляется в форме (9) с условием $\tau_0 < q \leqslant \tau$. Обозначая символом $I_0(N)$ часть интеграла $I(N)$, отвечающую основным интервалам, и символом $I_1(N)$ — часть, отвечающую дополнительным интервалам, будем иметь

$$I(N) = I_0(N) + I_1(N).$$

Из $I_0(N)$ выделяем главный член асимптотической формулы. Наконец, доказываем, что $I_1(N)$ является величиною порядка (относительно N) низшего, чем порядок главного члена; для этой последней цели и необходимы оценки сумм (6).

Примерно теми же средствами, как и проблема Варинга (метод Харди и Литльвуда и оценки сумм (6) помощью метода Вейля или же моего метода 1934 г.), были решены и другие важные проблемы, которые могут называться *обобщениями проблемы* Варинга; так, например, была всесторонне исследована проблема о представлении целого числа N в форме

$$N = \psi(x_1) + \ldots + \psi(x_r),$$

где $\psi(x)$ — целый многочлен степени n, принимающий для целочисленных значений x целочисленные значения. Большинство таких проблем является частными случаями общей проблемы о представлении целого числа N в форме

$$N = f(x_1, \ldots, x_r), \tag{10}$$

где $f(x_1, \ldots, x_r)$ — целая рациональная функция, принимающая для целочисленных значений $x_1, \ldots, x_r$ целочисленные значения; при этом характерной особенностью всех таких проблем является то, что совокупность Ω состоит из всех без исключения целых точек $(x_1, \ldots, x_r)$ некоторой r-мерной области.

Однако проблему о представлении целого числа N в форме (10) можно ставить и при условии, что совокупность Ω состоит не из всех целых точек r-мерной области, а только из части этих точек. Мы остановимся далее лишь на случае, когда совокупность Ω состоит только из тех целых точек r-мерной области, у которых при $0 < s \leqslant r$ первые s координат $x_1, \ldots, x_s$ суть простые числа. Это — наиболее важный класс „аддитивных проблем теории простых чисел“. Для удобства дальнейшего изложения при $s < r$ проблему будем называть *смешанной*, а при $s = r$ — *чистой*. Главной целью настоящей статьи является рассмотрение „чистых“ проблем. Еще в 1742 г. из переписки Гольдбаха с Эйлером возникла так называемая *проблема Гольдбаха* — гипотеза, согласно которой всякое целое число, превосходящее 2, есть сумма не более трех нечетных простых чисел. В частности, согласно этой гипотезе, четное число, превосходящее 2, должно быть суммою двух простых чисел. Пытаясь решить последнюю задачу, Брун (в 1919 г.) разработал особый элементарный метод, в основу которого он положил общеизвестную идею *решета Эратосфена*. Помощью этого метода Брун показал, что всякое положительное четное число есть сумма двух слагаемых, каждое из которых является произведением не более чем 9 простых чисел. В дальнейшем число 9 было заменено числом 4, но доказать таким путем гипотезу Гольдбаха для четных чисел все же не удалось. Шнирельман, присоединив к методу Бруна свои соображения о плотности последовательности, состоящей из целых положительных чисел, показал (в 1930 г.), что всякое целое число, превосходящее 1, есть сумма ограниченного числа r простых чисел; в дальнейшем было пока-

зано, что r не превосходит 67. Харди и Литльвуд (в 1920 г.) применили к решению проблемы Гольдбаха для нечетных чисел свой метод. Они условно вывели асимптотическую формулу для числа $I(N)$ представлений нечетного числа N в форме

$$N = p_1 + p_2 + p_3,$$

где p_1, p_2, p_3 — простые числа; из этой асимптотической формулы справедливость гипотезы Гольдбаха для нечетных чисел следовала уже тривиально. Условность вывода Харди и Литльвуда заключалась в применении некоторых недоказанных теорем, относящихся к теории так называемых *L-рядов*. Если принять как основу схему вывода асимптотической формулы для $I(N)$, описанную нами выше, то эта условность распространялась как на часть вывода, касающуюся выделения из $I_0(N)$ главного члена асимптотической формулы, так и на часть вывода, касающуюся оценки $I_1(N)$. Впоследствии (еще до 1937 г.) на основе одной формулы Пейджа (1935 г.), выражающей приближенно число простых чисел арифметической прогрессии, не превосходящих данного большого числа, были разработаны методы (метод, основанный на известной лемме Зигеля, и метод Эстермана, дающий несравнимо более грубую оценку порядка остаточного члена, но не зависящий от леммы Зигеля), позволяющие часть вывода, касающуюся $I_0(N)$, проводить вполне строго как для смешанных, так и для чистых проблем. А так как для смешанных проблем оценка $I_1(N)$ может осуществляться примерно теми же средствами, что и для упомянутых выше обобщений проблемы Варинга (оценки сумм (6)), то указанное обстоятельство открыло широкий путь к решению самых разнообразных смешанных проблем. Так, например, Эстерманом уже к началу 1937 г. была доказана представимость достаточно большого целого числа N в форме

$$N = p_1 + p_2 + x^2,$$

где p_1 и p_2 — простые числа и x — целое положительное число. Что же касается чистых проблем, то здесь оценка $I_1(N)$ требовала уменья нетривиально оценивать суммы, получаемые из сумм (6) введением ограничения, что $x_1, \ldots, x_k$ пробегают только простые числа. В 1937 г. мною было показано, что последняя задача может быть успешно разрешена применением моего метода 1934 г. Это обстоятельство открыло,

наконец, широкий путь и к решению самых разнообразных чистых проблем.

Первою была доказана асимптотическая формула для числа $I(N)$ представлений положительного нечетного числа N в форме

$$N = p_1 + p_2 + p_3,$$

где p_1, p_2, p_3 — простые нечетные числа. Эта формула такова:

$$I(N) = \frac{N^2}{2(\log N)^3} S(N) + O\left(\frac{N^2}{(\log N)^{3,5-\varepsilon}}\right), \tag{11}$$

где ε — произвольно малое положительное постоянное и

$$S(N) = \prod_p \left(1 - \frac{1}{(p-1)^2}\right) \prod_p{}' \left(1 - \frac{1}{p^2 - 3p + 3}\right),$$

причем $\prod\limits_p$ распространяется на все простые числа, а $\prod\limits_p{}'$ — лишь на простые числа, делящие N. Из этой асимптотической формулы справедливость гипотезы Гольдбаха для достаточно больших нечетных чисел получается уже как весьма частное следствие.

Была исчерпывающе исследована (ван-дер-Корпут) проблема о представлении целого числа N в форме

$$N = K_1 p_1 + K_2 p_2 + K_3 p_3,$$

где K_1, K_2, K_3 — заданные целые числа и p_1, p_2, p_3 — простые нечетные числа.

Было показано (независимо друг от друга пятью авторами), что гипотеза Гольдбаха о представимости четных чисел суммами двух простых нечетных чисел справедлива для «почти всех» четных чисел, а именно: отношение к N числа таких четных чисел, не превосходящих N, которые суммою двух нечетных простых чисел не представимы, стремится к нулю при неограниченном возрастании N. (В частности, применением леммы Зигеля доказано, что это отношение является величиною порядка $(\log N)^{-h}$, где h — произвольно большое положительное постоянное.) Однако отсюда до полного доказательства гипотезы Гольдбаха для четных чисел еще очень далеко, и эта задача еще ждет своего решения. Зато из указанных выше результатов, касающихся

представления нечетного числа суммою трех нечетных простых чисел, сразу следует, что всякое достаточно большое четное число представляется суммою четырех нечетных простых чисел, причем легко получается и асимптотическая формула для числа представлений.

Была выведена асимптотическая формула для числа $I(N)$ представлений целого положительного числа N в форме

$$N = p^n{}_1 + \ldots + p^n{}_r, \tag{12}$$

где $p_1, \ldots, p_r$ — нечетные простые числа, n — целое число, превосходящее 1 и $r > r_0(n)$, причем $r_0(n)$ — некоторая функция, при $n \geqslant 12$ равная $[10n^2 \log n]$. Эта формула такова:

$$I(N) = \frac{(\Gamma(1+\nu))^r}{\Gamma(r\nu)} \frac{N^{r\nu-1}}{l^r} S_n(N) + O\left(\frac{N^{r\nu-1}}{l^{r+\delta}}\right). \tag{13}$$

Здесь $\nu = n^{-1}$, $l = \nu \log N$, δ — положительное постоянное число,

$$S_n(N) = \sum_{q=1}^{\infty} W(q), \qquad W(q) = (\varphi(q))^{-r} \sum_a{}' R^r{}_{a,q} e^{-2\pi i \frac{a}{q} N},$$

$$R_{a,q} = \sum_x{}' e^{2\pi i \frac{a}{q} x^n},$$

причем символом $\sum\limits_s{}'$ обозначается сумма, распространенная на значения s, взаимно простые с q с условием $0 \leqslant s < q$. Весьма правдоподобной является гипотеза, что функция $r_0(n)$ может быть заменена функцией cn, где c — положительное постоянное (не зависящее от n).

Было исследовано $S_n(N)$, входящее в асимптотическую формулу (13). Пусть при четном $n \geqslant 2$ имеем $K = \prod\limits_p p^\gamma$, где p пробегает лишь простые числа с условием, что $p-1$ делит n, причем, обозначая буквою h наибольшее целое число с условием, что p^h делит n, имеем $\gamma = h+2$ при $p=2$ и $\gamma = h+1$ при $p > 2$. Пусть при нечетном $n > 1$ имеем $K = 2$. Было показано (*Hua*, 1938), что при

$$r \geqslant 4n, \qquad r \equiv N \pmod K$$

справедливо неравенство $S_n(N) > c$, где c — положительное число, не зависящее от N. При этом при $n = 2$ неравенство $r \geqslant 4n$ здесь может быть заменено и более точным неравенством $r \geqslant 5$.

Так как при $n = 2$ имеем $r_0(2) = 5$, $K = 24$, то отсюда следует, что при

$$r \geqslant 5, \qquad r \equiv N \pmod{24}$$

всякое достаточно большое целое число N может быть представлено в форме

$$N = p_1^2 + \ldots + p_r^2$$

с простыми нечетными $p_1, \ldots, p_r$. Например, при $N \equiv 5 \pmod{24}$ всякое достаточно большое целое число N может быть представлено в форме

$$N = p_1^2 + p_2^2 + p_3^2 + p_4^2 + p_5^2$$

с простыми нечетными p_1, p_2, p_3, p_4, p_5.

При $n > 2$ было установлено (*Хуа*, 1938) существование такой функции $r_1(n)$, асимптотически близкой к $4n \log n$, что при

$$r \geqslant r_1(n), \qquad r \equiv N \pmod{K}$$

всякое достаточно большое целое число N может быть представлено в форме (12).

Было решено и множество других самых разнообразных проблем. Например (p_s везде обозначает нечетное простое число), была установлена представимость почти всех четных чисел разностью $p_1 - p_2$ (ван-дер-Корпут), представимость почти всех четных чисел суммою $p_1 + p_2^k$, где k — нечетное, превосходящее 1 (*Хуа*), представимость всех достаточно больших нечетных чисел суммою $p_1 + p_2 + p_3^k$, где k — нечетное, превосходящее 1 (*Хуа*), и т. д.

Чтобы составить более отчетливое представление о методе, каким достигается решение «чистых» аддитивных проблем, остановимся на кратком описании вывода асимптотической формулы (11). Число $I(N)$ представлений нечетного числа N в форме

$$N = p_1 + p_2 + p_3$$

мы получим, если, положив в формуле (7) $r = 3$, $f(p_1, p_2, p_3) = p_1 + p_2 + p_3$, совокупность Ω составим из тех целых точек (p_1, p_2, p_3)

трехмерного пространства, координаты которых суть простые числа, не превосходящие N:

$$I(N)=\int_0^1 \sum_{p_1\leqslant N}\sum_{p_2\leqslant N}\sum_{p_3\leqslant N} e^{2\pi i\alpha(p_1+p_2+p_3-N)}d\alpha=$$

$$=\int_0^1 S_\alpha^3 e^{-2\pi i\alpha N}d\alpha,\quad S_\alpha=\sum_{p\leqslant N} e^{2\pi i\alpha p}.$$

Полагая $l=\log N$, $\tau_0=l^3$, $\tau=Nl^{-14}$, мы интервал $0\leqslant\alpha\leqslant 1$ интегрирования заменим таким:

$$-\tau^{-1}\leqslant\alpha\leqslant 1-\tau^{-1}. \tag{14}$$

Этот последний интервал мы разбиваем на основные интервалы, включающие все α вида

$$\alpha=\frac{a}{q}+z,\qquad (a,q)=1,\ -\frac{1}{\tau}\leqslant z\leqslant\frac{1}{\tau},\quad 0<q\leqslant l^3 \tag{15}$$

и дополнительные интервалы, остающиеся после выделения из интервала (14) основных интервалов. Всякое α дополнительного интервала может быть представлено в форме

$$\alpha=\frac{a}{q}+z,\qquad (a,q)=1,\qquad -\frac{1}{q\tau}\leqslant z\leqslant\frac{1}{q\tau},\qquad l^3<q\leqslant\tau. \tag{16}$$

Соответственно указанному подразделению интервала интегрирования на основные и дополнительные интервалы интеграл $I(N)$ разобьется на сумму двух слагаемых

$$I(N)=I_0(N)+I_1(N).$$

Дальнейшие рассуждения будут основаны на двух следующих леммах.

Лемма 1 (следствие леммы Пейджа). Пусть при данном произвольно малом положительном ε_0 заданы произвольно большие c_1 и c. Тогда число $\pi(N,q,t)$ простых чисел, не превосходящих N и заключенных в арифметической прогрессии $qx+t$; $x=0, 1, 2, \ldots$, удовлетворяющей условиям

$$(q,t)=1,\qquad 0\leqslant t<q\leqslant l^{c_1},$$

выражается формулой

$$\pi(N,\ q,\ t) = \frac{1}{\varphi(q)} \int_{2}^{N} \frac{dx}{\log x} + H,$$

где для всех q, кроме, быть может, ряда исключительных q, являющихся кратными какого-то одного $q = q_0$, удовлетворяющего условию

$$q_0 \geqslant l^{2-\varepsilon_0},$$

при некотором постоянном c_2 имеем неравенство

$$|H| \leqslant c_2 \frac{Nl^{-c}}{l\,\varphi(q)}.$$

Лемма 2 (мой метод). Обозначая буквою ε_1 произвольно малое положительное постоянное, для α, принадлежащего основному интервалу (условие (15)), будем иметь

$$|S_\alpha| \leqslant c_3 \frac{Nl^{\varepsilon_1}}{l\sqrt{q}},$$

а для α, принадлежащего дополнительному интервалу (условие (16)), будем иметь

$$|S_\alpha| \leqslant c_3 N l^{-2.5+\varepsilon_1},$$

где c_3 — постоянное, зависящее только от ε_1.

Применяя лемму 1 и частично лемму 2, мы без особого труда выделим из $I_0(N)$ главный член асимптотической формулы (11) с точностью до величины порядка, указанного в обозначении остаточного члена той же формулы. Оценка $I_1(N)$ производится следующим образом. Имеем

$$I_1(N) = \int_{(1)} S_\alpha^3 e^{-2\pi i\alpha N} d\alpha,$$

где интегрирование распространяется лишь на дополнительные интервалы. Отсюда находим

$$|I_1(N)| \leqslant \int_{(1)} |S_\alpha^3 e^{-2\pi i\alpha N}|\, d\alpha \leqslant \int_{(1)} |S_\alpha|^3 d\alpha = \int_{(1)} |S_\alpha|\, |S_\alpha|^2\, d\alpha.$$

Поэтому, согласно лемме 2,

$$|I_1(N)| \leqslant \int_{(1)} c_3 N l^{-2.5+\varepsilon_1} |S_\alpha|^2 \, d\alpha \leqslant c_3 N l^{-2.5+\varepsilon_1} \int_{(1)} |S_\alpha|^2 \, d\alpha.$$

А так как дополнительные интервалы суть части интервала $0 \leqslant \alpha \leqslant 1$, то отсюда находим

$$|I_1(N)| < c_3 N l^{-2.5+\varepsilon_1} \int_0^1 |S_\alpha|^2 \, d\alpha.$$

Но $|S_\alpha|^2 = S_\alpha \overline{S_\alpha}$, где $\overline{S_\alpha}$ — комплексное число, сопряженное с S_α, получаемое заменою в выражении для S_α числа i на число $-i$; следовательно

$$|S_\alpha|^2 = \sum_{p_1 \leqslant N} e^{2\pi i \alpha p_1} \sum_{p_2 \leqslant N} e^{-2\pi i \alpha p_2} = \sum_{p_1 \leqslant N} \sum_{p_2 \leqslant N} e^{2\pi i \alpha (p_1 - p_2)},$$

$$\int_0^1 |S_\alpha|^2 \, d\alpha = \sum_{p_1 \leqslant N} \sum_{p_2 \leqslant N} \int_0^1 e^{2\pi i \alpha (p_1 - p_2)} \, d\alpha.$$

А так как интеграл $\int_0^1 e^{2\pi i \alpha (p_1 - p_2)} \, d\alpha$ равен 1 при $p_1 = p_2$ и равен нулю в противном случае, то

$$\int_0^1 |S_\alpha|^2 \, d\alpha = \pi(N),$$

где $\pi(N)$ — число простых чисел, не превосходящих N. Поэтому, пользуясь известным неравенством $\pi(N) < 2Nl^{-1}$, из всего доказанного можем заключить, что

$$I_1(N) \leqslant c_3 N l^{-2.5+\varepsilon_1} \, 2Nl^{-1} = 2c_3 N^2 l^{-3.5+\varepsilon_1},$$

что и докажет полностью формулу (11).

Член-корреспондент АН СССР

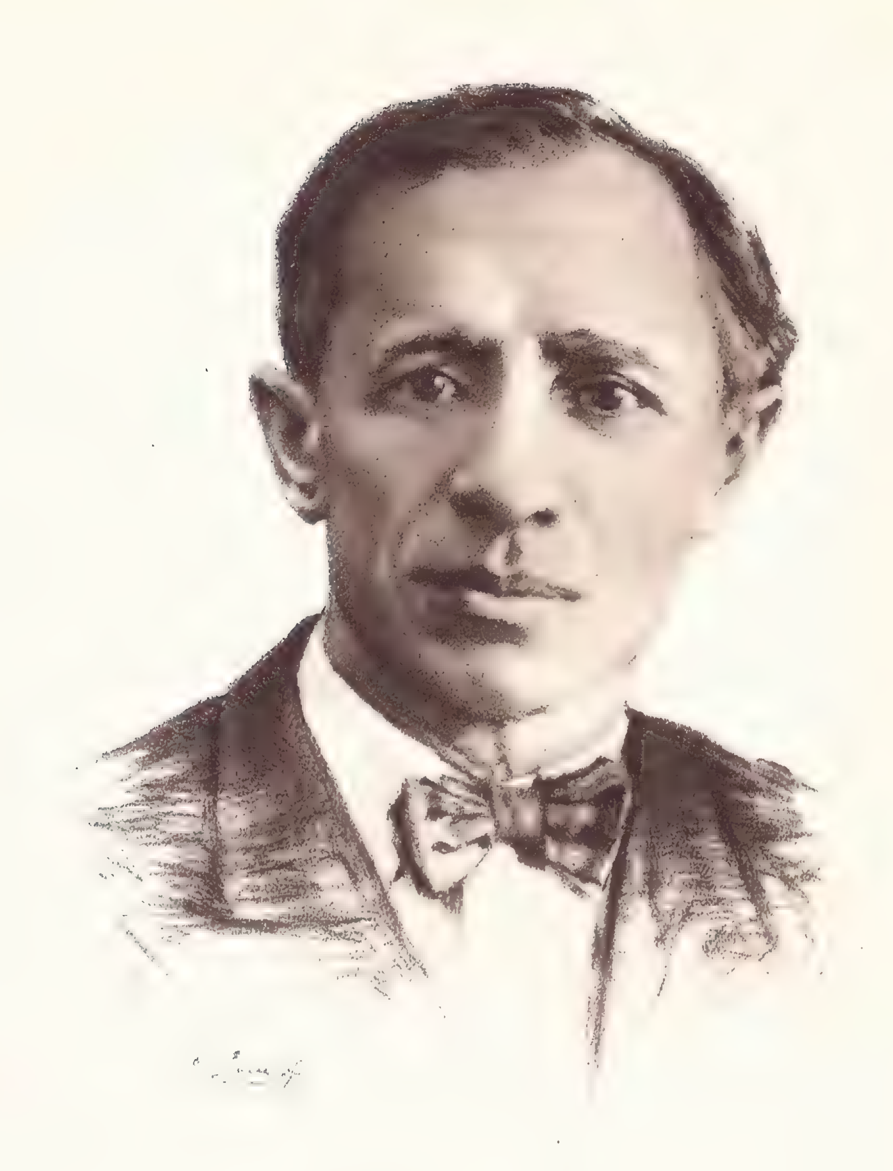
Н. Г. Чеботарев

ПРОБЛЕМА РЕЗОЛЬВЕНТ

В истории математики известно немало случаев, когда та или иная проблема привлекала внимание ученых из эстетических или чисто теоретических побуждений, и только по истечении многих лет и даже веков обнаруживалось ее практическое значение. Кроме общеизвестного примера теории вероятностей, можно назвать цикл задач на построение при помощи циркуля и линейки. Нас может удивить, с какой настойчивостью древние греки решали делосскую задачу о нахождении стороны куба с удвоенным объемом, стремясь дать теоретическое решение при помощи циркуля и линейки для задачи, которую на практике можно было бы решить точнее, поскольку циркуль — инструмент, всегда приводящий к неточностям.

У греков этот интерес к теоретическим решениям, возможно, был связан с их пренебрежительным отношением к практическим задачам. Но как нам ни чужд образ мыслей древних греков, он сыграл в истории науки большую положительную роль. Не говоря уже о развитой ими строгости логической мысли, давшей возможность впоследствии поднять на огромную высоту уровень точных наук, самая постановка теоретически точных задач позволила четко расклассифицировать задачи, увидеть, какая из них решается какими средствами.

В XIX веке возникли доказательства невозможности решения опре-



деленных задач определенными средствами: деления круга на три части циркулем и линейкой, представления некоторых интегралов через элементарные функции и т. п. Этот круг задач, в свою очередь, оказался полезным для практики: он дал возможность экономить труд при решении так называемых серийных задач, на которых в дальнейшем остановимся подробнее.

Вернемся к задачам на построение циркулем и линейкой. Невозможность решения некоторых из них была доказана при помощи теории алгебраического решения уравнений, созданной молодым гениальным математиком Эваристом Галуа [1]. Незадолго до появления его работ Абель [7] доказал невозможность решения в радикалах уравнений выше 4-й степени. Галуа же дал способ узнавать относительно каждого данного уравнения, решается ли оно в радикалах или нет. Для этого он сопоставил с каждым уравнением группу (получившую в последнее время название группы Галуа), т. е. совокупность подстановок, которые можно производить над его корнями без нарушения существующих между ними рациональных соотношений. Оказалось, что вопрос о возможности решить заданное уравнение в радикалах вполне определяется структурой его группы Галуа. Не будем описывать тех свойств, которые присущи разрешимым группам, т. е. группам, которым соответствуют разрешимые в радикалах уравнения.

Теория Галуа дала возможность также ответить на вопрос, возможно ли выполнить всякую заданную задачу на построение при помощи циркуля и линейки. Для этого нужно построить уравнение, от которого зависит искомая для задачи величина, и найти группу этого уравнения. Чтобы решение было возможно, необходимо и достаточно, чтобы число подстановок, содержащихся в группе, было степенью двойки. Идя таким путем, Гаусс еще до Галуа дал «теоретически точное» построение правильного 17-угольника [4]. Таким же путем была доказана невозможность деления угла на три равные части, а также построения стороны куба с удвоенным объемом при помощи циркуля и линейки.

Введенное Галуа понятие группы играет в современной математике роль, далеко выходящую за пределы задачи решения уравнений в радикалах. Наряду с конечными группами подстановок, операции которых состоят в перестановке некоторого конечного числа предметов, были введены непрерывные группы, теорию которых создал Ли [5]. Их

операции состоят из определенного типа перемещений точек в пространстве. Для лучшего уяснения этого понятия возьмем гидродинамическую аналогию. Представим себе пространство заполненным жидкостью (или газом), находящейся в непрерывном движении. Запомним положение каждой ее частицы в какой-нибудь определенный момент времени t_0. В каждый момент времени t положение каждой частицы будет другим. Будем говорить, что расположение частиц в момент t переводится из их расположения в момент t_0 преобразованием нашей непрерывной группы, которое внутри группы определяется для каждого момента t. При этом движение жидкости должно быть установившимся. Это значит, что, совершая подряд два любых преобразования из одной и той же группы, мы опять придем к преобразованию той же группы.

Переменная t, значения которой определяют преобразования внутри группы, называется параметром группы. Наша аналогия предусматривает случай группы с одним параметром; обыкновенно же рассматриваются группы с несколькими параметрами. Так, одна из наиболее простых групп, преобразования которой определяются перемещениями в пространстве твердого тела, есть группа с шестью параметрами, поскольку положение твердого тела в пространстве определяется шестью величинами.

Вернемся к вопросу о решении уравнений в радикалах. Казалось бы, что этот вопрос потерял всякое практическое значение, поскольку известно, что не всякое уравнение решается в радикалах, а узнать, решается ли оно, — далеко не легкая задача. Гораздо проще применить один из хорошо разработанных в настоящее время способов численного решения уравнений, позволяющих находить корни уравнений любой степени с любой точностью. Все это совершенно справедливо до тех пор, пока мы имеем дело с одним уравнением или пока нам изредка приходится решать разрозненные уравнения. Но современная наука и техника столкнулись с необходимостью решать громадное число уравнений более или менее однородного типа, но с различными числовыми значениями входящих в него параметров. Лучшим примером этому может служить астрономия. Для вычисления элементов орбит по нескольким наблюдениям астрономам приходится решать уравнения нескольких определенных типов. Так, для вычисления орбит комет (с параболическими орбитами) по трем наблюдениям нужно решать кубические уравнения. Для

вычисления орбит планет нужно решать определенного типа уравнения 8-й степени с входящими в выражения коэффициентов данными наблюдений в качестве параметров. При громадном числе комет и планет (астероидов) астрономам приходится вычислять корни очень большого числа уравнений одинакового типа. Очевидно, что составление таблиц для такого рода уравнений значительно сократит их труд. Для кубических уравнений астрономы и пользуются особыми таблицами Баркера. Для уравнений же высших степеней составление таблиц тормозится большим числом входящих в коэффициенты параметров.

Очень просто составить таблицы каких-нибудь функций от одного аргумента; подавляющее большинство существующих таблиц принадлежит именно к этому типу. Таблицы функций от двух аргументов очень неудобны для пользования. Это знает всякий, пользовавшийся большими таблицами умножения.

А вот другой пример. Астрономы (равно как и другие вычислители) весьма заинтересованы в том, чтобы по данным $\log a$ и $\log b$ быстро и просто находить значение $\log (a+b)$. Но таблица от двух аргументов, дающая эти значения, была бы очень неудобна для пользования. Вычислители вменяют в большую заслугу Гауссу изобретение принципа, в силу которого эта задача решается при помощи таблицы от одного аргумента.

Этот принцип крайне прост [2, 3]: таблица содержит значения величин $\log\left(1+\frac{b}{a}\right)$ по данным значениям $\log a - \log b = \log \frac{a}{b}$.

Для нахождения значений функции от двух или даже большего числа аргументов очень полезны методы номографии. Но, чем больше аргументов, тем сложнее номограмма, тем труднее пользование ею и тем менее надежен результат. Таким образом, задача максимального уменьшения числа параметров, входящих в коэффициенты уравнения, приобретает большую важность.

Если бы всякое уравнение решалось в радикалах, то упомянутая задача получила бы идеально простое решение. В самом деле, каждый радикал есть функция только одного аргумента (если значения последнего комплексны, то получится функция от двух аргументов; но, представив аргумент в тригонометрической форме, мы приведем радикал к функциям от одного аргумента). Если известно выражение корней

6*

уравнения через радикалы, то для вычисления этих корней достаточно иметь набор таблиц радикалов соответствующих степеней и производить над получаемыми из них данными вычисления в пределах четырех арифметических действий.

К сожалению, таких радикальных выражений в природе не существует. Но их значение для «серийных задач» наводит на такую мысль: какой смысл искать выражения корней уравнения в радикалах, т. е. в форме выражений через функции одного аргумента весьма частного вида, когда мы имеем гораздо больше шансов на успех, если станем отыскивать выражения для корней в форме многократно повторяемых функций одного аргумента, не задаваясь заранее видом этих последних?

Такое обобщение задачи вначале имело успех. Оказалось, что уравнения 5-й степени, которые, как известно, не решаются в радикалах, тем не менее допускают представление своих корней через функции от одного аргумента. Это в принципе было известно очень давно. Чтобы разъяснить это, станем на несколько другую точку зрения. Если

$$f(x)=0 \tag{1}$$

заданное уравнение, а

$$y=\varphi(x) \tag{2}$$

любая рациональная функция с рациональными коэффициентами, то y удовлетворяет уравнению

$$F(y)=0 \tag{3}$$

той же степени. Это преобразование уравнения носит название преобразования Чирнгаузена. Если нам удастся подобрать функцию (2) так, чтобы соответствующее ей уравнение (3) содержало в коэффициентах всего один параметр, то задача будет решена. В самом деле, из равенства (2) определенным образом можно получить рациональное выражение корней x через корни y:

$$x=\psi(y). \tag{4}$$

С другой стороны, корень y уравнения (3) зависит только от одного параметра, входящего в коэффициенты. Из (4) следует, что x выражается рационально через функцию одного параметра.

Давно известно, что можно подобрать преобразование (2) так, чтобы самое общее уравнение 5-й степени

$$f(x)=x^5+a_1x^4+a_2x^3+a_3x^2+a_4x+a_5=0 \tag{5}$$

преобразовывалось в уравнение так называемого Бринг-Жерраровского вида

$$F(y)=y^5+py+q=0. \tag{6}$$

При этом коэффициенты функции (2) могут быть не рациональными, а быть корнями уравнений 2, 3 и 4-й степеней. Но поскольку все такие уравнения решаются в радикалах, они, следовательно, тоже выражаются через функции от одного аргумента. С другой стороны, в коэффициенты уравнения (6) входят два параметра p, q; но подстановка

$$y=\sqrt[5]{q}\cdot z$$

приводит уравнение (6) к виду

$$z^5+\frac{p}{\sqrt[5]{q^4}}z+1=0;$$

в последнем уравнении выражение $\frac{p}{\sqrt[5]{q^4}}$ является единственным параметром.

Можно дать более простое приведение уравнения (5) к однопараметрическому виду, если не ограничивать себя Бринг-Жерраровской формой. Именно, можно добиться того, чтобы в коэффициенты функции (2) входили только квадратные радикалы. Это сделал впервые Гальфен [8], пользуясь уравнениями деления аргумента эллиптической функции на 5. В алгебраической форме эти же результаты были получены Клейном [11, 15]; его уравнение (3) имеет вид

$$y^5+15y^4-10\gamma y^2+3\gamma^2=0, \tag{7}$$

где γ — параметр.

Этот результат Клейна связан с очень интересными соображениями из области теории групп, конечных и непрерывных. Группа Галуа уравнения (5) общего вида есть совокупность всех подстановок его пяти корней; таких подстановок всего $1\cdot2\cdot3\cdot4\cdot5 = 120$. Если же мы присоеди-

ним к области рациональности (т. е. условно будем считать рациональным) квадратный корень из так называемого дискриминанта уравнения (5), то группа Галуа понизится: в ней останутся только четные подстановки, всего числом 60. Эта группа — наименьшая из неразрешимых групп; она изоморфна (т. е. одинакова по структуре) группе вращений самого сложного правильного многогранника — икосаэдра. Представим себе икосаэдр вписанным в шар и будем считать вращением икосаэдра такой поворот шара вокруг центра, при котором каждая вершина икосаэдра попадет в положение какой-нибудь другой из его вершин (или останется на месте). Подсчитаем число различных вращений, учитывая, что икосаэдр имеет 12 вершин, составляющих шесть пар, каждая из которых лежит на одном и том же диаметре шара, и что вокруг каждой вершины расположено пять треугольных граней. Можно придать какой-нибудь определенной оси одно из шести положений; выбрав определенное положение оси, можно поменять местами обе находящиеся на ней вершины; наконец, фиксируя обе вершины, можно вращать вокруг них фигуру, придавая ей пять различных положений. Всего получится $6 \cdot 2 \cdot 5 = 60$ различных вращений.

Группа икосаэдра есть наибольшая из конечных групп вращений шара, поскольку икосаэдр — самый сложный из правильных многогранников. Другими словами, группа икосаэдра — наибольшая из конечных подгрупп группы всевозможных вращений шара. Последняя является непрерывной группой и зависит от трех параметров. Она изоморфна группе всевозможных дробных линейных преобразований

$$y = \frac{ax+b}{cx+d}. \tag{8}$$

Группа же дробных линейных преобразований является наибольшей группой из существующих в пространстве одного измерения групп. Этот факт делал весьма вероятной следующую гипотезу.

Если уравнение (1) можно преобразовать в уравнение (3), в коэффициенты которого входит k параметров, то его группа Галуа изоморфна конечной подгруппе некоторой непрерывной группы преобразований точек k-мерного пространства. Справедливо и обратное.

Эта гипотеза в неявном виде содержится в работах Клейна и других математиков, работавших над проблемой резольвент, или,

как она называлась в ранней стадии своего развития, проблемой форм. В явном виде мне удалось доказать ее в 1931 г. [12] и привести в более совершенную форму в 1933 г. [13].

Однако результаты, полученные таким образом для проблемы Клейна, были весьма неутешительны. Дело в том, что Виман [16] доказал, что знакопеременная группа степени $n \geqslant 8$ не может быть представлена как однородная линейная группа менее чем от $n - 1$ переменных. С другой стороны, Картан в своей диссертации высказал предположение, что все подгруппы максимального числа параметров у простых непрерывных групп (о которых только и должна итти здесь речь) регулярны, т. е. принадлежат к типу подгрупп, допускающих простое перечисление геометрическим методом, который был открыт Киллингом и усовершенствован Картаном. Известны все типы простых непрерывных групп: кроме пяти исключительных групп, они состоят из однородных линейных групп, ортогональных и так называемых симплициальных групп от n переменных. Из них только ортогональные группы содержат подгруппы с числом параметров на $n - 2$ меньшим, чем самые группы; у групп остальных типов они на $n - 1$ меньше. Из этого следует, что ортогональные группы могут быть представлены в пространстве не меньшего числа измерений, чем $n - 2$. Таким образом, знакопеременная группа из $n \geqslant 8$ цифр является подгруппой непрерывной группы, представляемой в пространстве числа измерений не меньшего, чем $n - 3$. Из этого, в силу приведенного результата, следует, что уравнение n-й степени общего вида, группа Галуа которого после присоединения к области рациональности квадратного корня из дискриминанта является знакопеременной группой, имеет резольвенту не менее чем с $n - 3$ параметрами. При этом коэффициенты преобразования могут содержать иррациональности; но их резольвенты не могут содержать более $n - 3$ параметров.

Можно было бы еще надеяться, что гипотеза Картана не верна. Тогда можно было бы уменьшить число параметров в резольвенте более чем на три параметра. Однако в 1938 г. мне удалось доказать правильность гипотезы Картана [14]. Таким образом, в приведенной постановке Клейна польза от решения проблемы резольвент имеет очень ограниченный характер.

Вместе с тем Гильберт предложил другую, расширенную формулировку проблемы резольвент. Она состоит в том, что задается число S па-

раметров, которые должна содержать резольвента. При этом коэффициенты уравнения (2) могут не быть рациональными, но уравнения с рациональными коэффициентами, которым они удовлетворяют, тоже допускают резольвенты, содержащие не более S параметров. Коэффициенты преобразования корней этих уравнений в корни резольвент, если они иррациональны, являются корнями уравнений, резольвенты которых опять содержат не более S параметров, и т. д. Этот процесс должен содержать конечное число шагов. Ищется наименьшее значение S, удовлетворяющее этим условиям.

Гильберт опубликовал свою формулировку проблемы резольвент в 1900 г. как одну из своих знаменитых 23 задач, которые на много лет определили направление работы математиков [9]. Из них 13-я задача посвящена проблеме резольвент. Собственно говоря, Гильберт формулировал более частную задачу: доказать, что для общего уравнения 7-й степени не существует резольвенты с двумя параметрами. В 1926 г. он доказал [10], что для общего уравнения 9-й степени существует резольвента с четырьмя параметрами. Виман [17] получил более общий результат: общее уравнение степени $n \geqslant 10$ имеет резольвенту с $\leqslant (n-5)$ параметрами. Другие значения числа параметров в проблеме Гильберта указывают, что проблема резольвент в формулировке Гильберта существенно отличается от проблемы Клейна.

При решении своей проблемы Гильберт (и точно так же Виман) пользовался частными свойствами форм определенных степеней, а потому его методы не могут быть распространены на уравнения высших степеней. Кроме того, он не задавался целью доказать, что найденные им значения для S являются наименьшими из возможных.

Таким образом, для решения проблемы резольвент Гильберта недоставало общего принципа. Этот принцип невозможно было извлечь из сложной формулировки Гильберта. Мне удалось найти его в 1943 г. [6], поставив более общую проблему резольвент. Для разъяснения сущности этой проблемы пришлось ввести понятие *группы монодромии* уравнения, в коэффициенты которого входит некоторое число, скажем m, параметров. Пусть эти параметры будут комплексными числами; будем считать их вещественные и мнимые части декартовыми координатами $2m$-мерного пространства, каждая точка которого будет, таким образом, соответствовать системе численных значений параметров. Поэтому каж-

дой точке пространства будет соответствовать n корней нашего уравнения. Если мы будем непрерывно двигать точку, то корни тоже будут непрерывно менять значения. Пусть теперь точка описала в пространстве некоторый замкнутый путь. Поскольку она вернулась в исходное положение, корни в своей совокупности вернутся на свои старые места. Это, однако, не значит, что каждый из корней вернется в свое старое положение. В общем случае корни претерпят подстановку. Совокупность подстановок, испытываемых корнями при пробеге точкой всевозможных замкнутых путей в пространстве, носит название группы монодромии уравнения. Можно показать, что группа монодромии есть группа Галуа уравнения, если в качестве области рациональности взять поле рациональных функций от параметров. Если уравнение неприводимо, то его группа монодромии транзитивна. Это значит, что, задав любые два корня, соответствующие заданной точке пространства, можно найти такой замкнутый путь, проходящий через эту точку, при обходе которого один заданный корень переместится в другой.

В теории аналитических функций доказана так называемая теорема монодромии, обычно приводимая для одной независимой переменной (т. е. для плоскости); но ее без труда можно распространить на любое число независимых переменных. Она состоит в том, что если вообще существуют замкнутые пути, вдоль которых корни уравнения перемещаются, то существуют точки, в любой окрестности которых содержатся замкнутые пути, вдоль которых корни испытывают перемещения. Такого рода точки называются критическими. По аналогии с теорией алгебраических чисел, будем называть группой инерции критической точки совокупность подстановок, которые испытывают корни уравнения при обходе точки по всевозможным замкнутым путям, расположенным в окрестности критической точки. Тогда теорему монодромии можно высказать в следующей расширенной формулировке:

Наименьшая группа подстановок, содержащая как подгруппы все группы инерции, соответствующие всевозможным критическим точкам $2m$-мерного пространства, есть группа монодромии.

В силу непрерывности корней уравнения как функций точки пространства при обходе точкой бесконечно малых замкнутых путей корни могут перемещаться только тогда, когда они бесконечно близки. Отсюда следует, что в критических точках некоторые из корней уравнения должны

сливаться. Таким образом, мы получим все многообразие критических точек, если приравняем нулю дискриминант уравнения как функцию параметров

$$D(\alpha_1, \alpha_2, \ldots, \alpha_m) = 0. \tag{9}$$

Не исключена возможность, что некоторые из точек, лежащих на многообразии (9), не будут критическими.

Для наших целей необходимо более тонкое различение критических точек. Если в какой-нибудь критической точке сольется только два корня уравнения, то единственным возможным перемещением корней при обходе вблизи этой критической точки будет транспозиция, т. е. перемещение двух близких корней. Это будет наиболее простая из возможных критических точек. Если же в критической точке сольются три корня или две различные пары корней, то при обходе вблизи нее станут возможными различные типы подстановок между корнями. Чем больше корней сливается в критической точке, тем сложнее сама критическая точка, а также соответствующая ей группа инерции. Для решения проблемы резольвент необходимо знать число различных типов критических точек, входящих в критическое многообразие. Для различения этих типов нами был предложен следующий прием. Пусть корни уравнения будут $x_1, x_2, \ldots, x_n$. Составим выражение

$$\prod_S (t_1 x_{\lambda_1} + t_2 x_{\lambda_2} + \ldots + t_n x_{\lambda_n}),$$

в котором будем считать $t_1, t_2, \ldots, t_n$ независимыми переменными, а произведение распространим на все подстановки

$$S = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \lambda_1 & \lambda_2 & \lambda_3 & \ldots & \lambda_n \end{pmatrix}$$

группы монодромии уравнения. Получится форма (однородный полином) от $t_1, t_2, \ldots, t_n$, коэффициенты которой, как не меняющиеся от подстановок группы монодромии, рационально выражаются через коэффициенты уравнения и, следовательно, через параметры $\alpha_1, \alpha_2, \ldots, \alpha_m$. Обозначим эту форму через

$$\Phi(t_1, t_2, \ldots, t_n). \tag{10}$$

Подставим в нее значения параметров $\alpha_1, \alpha_2, \ldots, \alpha_m$, соответствующие испытуемой критической точке. Чтобы в ней имели место следующие совпадения корней

$$\begin{aligned} x_1 = x_2 = \ldots = x_{\mu_1}, \\ x_{\mu_1+1} = \ldots = x_{\mu_2}, \\ \ldots\ldots\ldots \\ x_{\mu_{k-1}+1} = \ldots = x_n, \end{aligned} \tag{11}$$

необходимо и достаточно, чтобы форма (10) обращалась в нуль, если положить

$$\begin{aligned} t_1 + t_2 + \ldots + t_{\mu_1} = 0, \\ t_{\mu_1+1} + \ldots + t_{\mu_2} = 0, \\ \ldots\ldots\ldots \\ t_{\mu_{k-1}+1} + \ldots + t_n = 0. \end{aligned} \tag{12}$$

При этом мы не можем различать равенств (11) от равенств, получаемых из (11) применением к $x_1, x_2, \ldots, x_n$ любой подстановки группы монодромии. Таким образом, если ввести обозначение

$$S = (x_1, x_2, \ldots, x_{\mu_1})(x_{\mu_1+1}, \ldots, x_{\mu_2}) \ldots (x_{\mu_{k-1}+1}, \ldots, x_n),$$

то в случае выполнения равенств (11) мы будем говорить, что наша критическая точка соответствует подстановке S или более высокой подстановке, если, кроме (11), в ней имеются и другие слияния корней. Ограничимся случаем, когда параметры $\alpha_1, \alpha_2, \ldots, \alpha_m$ входят в коэффициенты уравнения линейно. Тогда если критическая точка соответствует подстановке S, то в ее группу инерции входит подстановка S (или подобные ей), а также более низкие подстановки, т. е. подстановки, которым соответствуют равенства, являющиеся частью равенств (11).

Таким образом, чтобы узнать, соответствует ли данная критическая точка подстановке S (или более высокой), надо к форме Φ применить равенства (12) (например, выражая $t_{\mu_1}, t_{\mu_2}, \ldots, t_n$ через остальные переменные) и проверить, все ли коэффициенты полученной таким образом формы Φ_S обращаются в этой точке в нуль.

Чтобы выяснить, на какие категории подразделяются критические точки в данном уравнении, расположим входящие в нее подстановки по высоте

$$S_1 \subset S_2 \subset \ldots \subset S_q.$$

При этом может случиться, что такой цепочкой не исчерпаются все подстановки группы монодромии. Тогда надо составить всевозможные цепочки.

Приравняем нулю коэффициенты формы Φ_{S_1}; получим многообразие из самых низших критических точек (относительно данной цепочки), а также из высших, соответствующих подстановкам $S_2, \ldots, S_q$. Затем приравняем нулю коэффициенты формы Φ_{S_2}. Может случиться, что мы не получим новых соотношений между параметрами; это будет означать, что все точки, соответствующие S_1, соответствуют более высокой подстановке S_2 или еще выше. Продолжая делать то же относительно форм $\Phi_{S_3}, \ldots, \Phi_{S_q}$, мы получим ряд критических многообразий, из которых каждое последующее будет содержаться в предыдущем. Каждое из последующих критических многообразий определяется, как это можно доказать, одним дополнительным уравнением между комплексными значениями параметров, т. е. двумя уравнениями между их вещественными и мнимыми частями. Из этого следует, что последовательные критические многообразия, соответствующие подстановкам одной цепочки, имеют соответственно измерения $2m-2, 2m-4, \ldots, 2m-2q_1$, где $q_1 \leqslant q$.

Теперь можно формулировать проблему резольвент в более общем и, как нам кажется, в более естественном виде. Пусть даны два уравнения одной и той же степени и пусть коэффициенты второго из них суть функции коэффициентов первого. Тогда замкнутому пути в пространстве, соответствующем первому уравнению, будет соответствовать замкнутый путь в пространстве, соответствующем второму уравнению. Таким образом, подстановки групп монодромии обоих уравнений мы приведем в соответствие, носящее характер изоморфизма: произведению двух подстановок одной группы соответствует произведение подстановок соответствующих множителей второй группы. В частности, критической точке соответствует критическая точка, причем, ввиду соответствия замкнутых путей в их окрестностях, будут подобны их группы инерции, в силу чего обе критические точки должны быть одной и той же высоты. Другими словами, в обоих пространствах устанавливаются цепочки равной длины подстановок и критических многообразий, и тогда критической точке, соответствующей определенному номеру подстановки, в другом пространстве соответствует критическая точка, соответствующая подстановке с тем же номером. При этом учтем, что пространства для обоих уравнений

могут быть различных измерений. Однако из того, что в первом уравнении существует q_1 различных категорий критических точек, следует, что во втором уравнении их должно быть также q_1. Поэтому, если мы обозначим через m_2 число параметров в коэффициентах второго уравнения, то

$$2m_2 - 2q_1 \geqslant 0,$$

откуда

$$m_2 \geqslant q_1. \tag{13}$$

Итак, под резольвентой заданного уравнения мы будем разуметь уравнение, коэффициенты которого можно поставить в зависимость от коэффициентов заданного уравнения так, чтобы отдельные корни одного из уравнений были однозначными аналитическими функциями отдельных корней другого уравнения, т. е. чтобы при обходе замкнутых путей в пространствах этих уравнений корни уравнений, при надлежащей нумерации, испытывали те же подстановки. При этом может случиться, что замкнутым путям в одном пространстве будут соответствовать разомкнутые пути в другом.

Этим не исчерпываются требования, которые мы налагаем на резольвенту: резольвента должна еще содержать в своих коэффициентах возможно меньшее число параметров. Из наших рассуждений относительно критических многообразий вытекает, что число параметров в коэффициентах резольвенты не может быть меньше, чем максимальная длина цепочки подстановок в группе монодромии уравнения, причем, конечно, мы должны выкинуть из цепочки те подстановки, которым не соответствуют критические многообразия, так что это число в наших обозначениях есть q_1, а не q. Можно ли действительно построить резольвенту с q_1 параметрами для в с я к о г о заданного уравнения, до настоящего времени не удалось решить. Кроме того, до сих пор не выяснен вопрос, что изменится в полученных результатах, если освободиться от требования, чтобы параметры входили в коэффициенты линейно.

Рассмотрим в виде примера общее уравнение n-й степени, в котором будем полагать старший коэффициент равным единице, а остальные считать независимыми переменными. Его группа монодромии есть симметрическая группа; но, присоединив к области рациональности квадратный корень из дискриминанта этого уравнения, мы снизим его группу монодромии до знакопеременной группы, которая, как известно, проста.

В знакопеременной группе содержатся следующие цепочки подстановок:

$$(123) \subset (12345) \subset (1234567) \subset \ldots \subset \left(123 \ldots 2\left[\frac{n}{2}\right] - (-1)^n\right),$$

где $2\left[\frac{n}{2}\right] - (-1)^n$ есть самое большое нечетное число, не превышающее n. Длина этой цепочки равна

$$\left[\frac{n}{2}\right] - \frac{1+(-1)^n}{2} = \left[\frac{n-1}{2}\right]. \qquad (14)$$

С другой стороны, в знакопеременной группе не содержится цепочек, длина которых превышает число (14). В самом деле, отмечая в каждой подстановке число содержащихся в ней циклов, включая в это число и одночленные циклы (т. е. инвариантные цифры), заметим, что для четной подстановки число имеет ту же четность, что и n. С другой стороны, если, например,

$$S_i \subset S_{i+1},$$

то подстановка S_{i+1} непременно содержит меньшее число циклов, чем S_i. В силу одинаковой четности эти числа должны отличаться по крайней мере на 2. Поэтому числа циклов в подстановках цепочки

$$S_k \supset S_{k-1} \supset \ldots \supset S_2 \supset S_1$$

не могут быть меньше, чем соответственно числа

$$1, 3, 5, \ldots, 2k-1.$$

Но S_1 не может быть ни тождественной подстановкой, ни транспозицией, в силу чего

$$2k-1 \leqslant n-2,$$

откуда

$$k \leqslant \left[\frac{n-1}{2}\right].$$

Таким образом, искомая нижняя граница для числа параметров резольвенты равна

$$S = \left[\frac{n-1}{2}\right]. \qquad (15)$$

Сопоставим эти значения с теми, которые были получены Гильбертом для $5 \leqslant n \leqslant 9$:

n	5	6	7	8	9
S [по формуле (15)]	2	2	3	3	4
S [по Гильберту]	1	2	3	4	4

Это сопоставление показывает, что при $n=5$ введение иррациональных резольвент на самом деле снижает число параметров. Случай же $n=8$ требует дополнительного исследования. Не исключена возможность, что примененный Гильбертом частный прием нахождения резольвенты не дает минимального числа параметров.

Предложенный способ нахождения числа параметров для резольвент дает перспективу к нахождению решения проблемы резольвент для произвольного n, причем не только для уравнений общего типа, но также для всякого конкретно заданного уравнения, у которого коэффициенты являются определенными функциями от параметров.

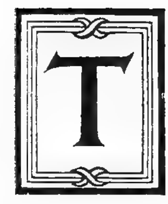
ЛИТЕРАТУРА

1. *Галуа.* Соч. ОНТИ, 1936.
2. *Гаусс.* Пятизначные логарифмические таблицы. ГТТИ, 1933.
3. *Пржевальский.* Пятизначные таблицы логарифмов.
4. *Чеботарев.* Основы теории Галуа. ОНТИ, 1937.
5. *Чеботарев.* Теория групп Ли. ОНТИ, 1940.
6. *Чеботарев.* Проблема резольвент и критические многообразия. Изв. АН СССР, серия матем., 1943, т. 7, стр. 123—126.
7. *Abel.* Démonstration de l'impossibilité de la résolution algébrique des équations générales qui passent le quatrième degré. J. f. reine angew. Math. 1826, Bd. 1; Oeuvres complètes. Christiania, 1881, p. 66—94.
8. *Halphen.* Traitè des fonctions elliptiques. T. 3.
9. *Hilbert.* Mathematische Probleme. Gött. Nachr. 1900, S. 253—297; Ges. Abh. 1935, Bd. 3, Berlin, S. 290—329.
10. *Hilbert.* Ueber die Gleichung neunten Grades. Math. Ann., Bd. 97, 1927, S. 243—250; Ges. Abh., 1933, Bd. 2, Berlin, S. 393—400.
11. *Klein.* Ges. Math. Abh., 1922, Bd. 2. Berlin, S. 255—504.
12. *Tschebotaröw.* Ueber ein algebraisches Problem von Herrn Hilbert. I. Math. Ann., Bd. 104, 1931, S. 459—471; II. Bd. 105, S. 240—255.
13. *Tschebotaröw.* Ueber das Klein-Hilbertsche Resolventenproblem. Bull. Soc. Math. de Kazan, 1933, Bd. 6, № 3, S. 5—22.
14. *Tschebotaröw.* Ueber irreguläre Darstellungen von halbeinfachen. Lieschen Gruppen. Comp. Math., 1938, Bd. 6, S. 103—117.
15. *Weber.* Lehrbuch der Algebra, 2. Braunschweig, 1899, S. 489.
16. *Wiman.* Ueber die Darstellung der symmetrischen und alternierenden Vertauschungsgruppen usw, Math. Ann., Bd. 52, 1899, S. 243—270.
17. *Wiman.* Ueber die Anwendung der Tschirnhausen-Transformation auf die Reduktion algebraischer Gleichungen. Nova Acta R. Soc. Sc. Uppsaliensis, vol. extra ordin. editum, 1927, S. 3—8.

Академик

М. А. Лаврентьев

ТЕОРИЯ КВАЗИКОНФОРМНЫХ ОТОБРАЖЕНИЙ



Теория квазиконформных отображений, развившаяся за последние 15 лет, представляет интерес как теория, широко обобщившая классическую теорию конформных отображений и вскрывшая геометрическую природу ряда фактов конформных отображений, получавшихся раньше только аналитически.

Особый интерес теория квазиконформных отображений представляет своими богатыми связями со многими узловыми проблемами анализа, геометрии и механики.

Замечу сейчас же, что теория квазиконформных отображений как теория, опирающаяся на качественные геометрические методы, не может претендовать в приложениях к дифференциальным уравнениям, в частности к задачам газовой динамики, на эффективные и численные решения этих задач и тем самым она не может заменить известных прямых методов и метода интегральных уравнений.

В настоящем кратком очерке теории квазиконформных отображений я постараюсь изложить сущность теории и ее связи главным образом с краевыми задачами систем дифференциальных уравнений с частными производными, в частности с задачами газовой динамики.

1. Предварительные понятия. Пусть в плоскости xOy задана область D и в этой области заданы две непрерывные и дифференци-

руемые функции

$$u=u(x,\ y), \qquad v=v(x,\ y) \tag{1}$$

или, в символике комплексных чисел,

$$z=x+iy, \qquad w=u+iv, \qquad i=\sqrt{-1}, \qquad w=f(z).$$

Пусть Δ есть совокупность точек плоскости uOv, соответствующих в силу (1) всем точкам области D. Говорят, что система (1) отображает область D на Δ. Отображение называется *однолистным*, если двум различным точкам D отвечают две различные точки Δ. При однолистном отображении, если D есть область, то Δ будет также областью.[1] Если отображение (1) области D на область Δ однолистно, то система (1), очевидно, может быть для всех точек u, v области Δ однозначно разрешена относительно x и y

$$x=x(u,\ v), \qquad y=y(u,\ v); \tag{2}$$

система (2) будет давать однолистное отображение Δ на D. Отображение (2) называется отображением, обратным отображению (1).

Пусть теперь x_0, y_0 есть произвольная точка области D, а u_0, v_0—ей соответствующая точка области Δ. В силу принятой дифференцируемости функций u, v мы можем построить главную линейную часть отображения (1) в окрестности точки x_0, y_0. Имеем

$$\begin{aligned} u-u_0&=u_x(x-x_0)+u_y(y-y_0),\\ v-v_0&=v_x(x-x_0)+v_y(y-y_0), \end{aligned} \tag{3}$$

где u_x, u_y, v_x, v_y — частные производные функций u, v, взятые в точке x_0, y_0. Функции u, v, вычисленные по (3), будут отклоняться от u, v, вычисленных по (1), на бесконечно малые высших порядков малости сравнительно с расстоянием ρ точки x, y до начальной точки x_0, y_0

$$\rho=\sqrt{(x-x_0)^2+(y-y_0)^2}.$$

Во всем дальнейшем мы будем предполагать, что всюду в области D определитель

$$\delta=u_xv_y-u_yv_x>0. \tag{4}$$

[1] В общем случае Δ может и выродиться, в частности может оказаться образом одного измерения — линией.

В силу этого для каждой фиксированной точки x_0, y_0 линейное преобразование (3) будет невырожденным и будет давать отображение всей плоскости xOy на всю плоскость uOv. Преобразование (3) будет переводить каждый круг плоскости xOy в некоторый эллипс плоскости uOv — порядок кривой всегда инвариантен относительно линейного преобразования. Пусть E есть эллипс плоскости xOy, переходящий в единичную окружность плоскости uOv. Обозначим через a большую ось E, через p отношение большой оси к малой оси, через t угол, образованный осью E с осью x, и, наконец, через α направление диаметра E, переходящего в прямую, параллельную оси u. Введенные параметры a, p, t и α с точностью до сдвига определяют преобразование (3) и могут быть выражены через u_x, u_y, v_x, v_y. В ряде случаев оказывается удобным вместо параметров a, p, t, α ввести другие параметры, так же характеризующиеся (3). Рассмотрим в плоскости u, v единичный квадрат с вершиной в точке $w_0=u_0+iv_0$, $i=\sqrt{-1}$ и со сторонами

$$\overline{w_0w_1},\quad \overline{w_0w_2},\quad w_2-w_0=(w_1-w_0)\,e^{i\frac{\pi}{2}}.$$

Обозначим через ν угол, образованный вектором $\overline{w_0w_1}$ с осью u

$$w_1-w_0=e^{i\nu}.$$

При отображении (3) построенный квадрат будет соответствовать некоторому параллелограму π_ν; пусть при этом точки w_1, w_2 соответствуют точкам z_1, z_2. Положим

$$z_1-z_0=V_\nu e^{i\alpha_\nu},\qquad \theta_\nu=\arg\frac{z_2-z_1}{z_1-z_0},\qquad W_\nu V_\nu\delta=1. \tag{5}$$

Введенные величины (при любом фиксированном ν) полностью определяют параллелограм π_ν и могут быть также элементарно выражены через коэффициенты преобразования (3). Заметим, что V_ν есть коэффициент растяжения в направлении, ν — длина основания π_ν, W_ν — высота π_ν, θ_ν — угол при вершине и α_ν — „угол поворота“.

Величины a, p, t, α, а также W, α, θ, V мы будем называть *характеристиками* отображения (3).

Вернемся к отображению (1). Пусть l есть кривая, которая при этом отображении переходит в окружность радиуса ρ с центром в точке $z_0=x_0+iy_0$; наряду с l построим эллипс e, подобный и подобно рас-

положенный эллипсу E. Если за коэффициент подобия принять ρ, то линии l и e будут отличаться на бесконечно малые высшего порядка малости сравнительно с ρ, так как любое однолистное дифференцируемое отображение переводит с точностью до малых высших порядков каждый бесконечно малый круг в бесконечно малый эллипс.

Рассмотрим условия, которым должно удовлетворять отображение, чтобы в каждой точке эллипс e был кругом. Для этого необходимо и достаточно, чтобы линейное преобразование (3) было *ортогональным* преобразованием, т. е. чтобы частные производные u и v удовлетворяли соотношениям

$$u_x = v_y, \qquad u_y = -v_x. \tag{6}$$

Система (6) есть классическая система уравнений Коши — Римана, а отображение (1) есть конформное отображение.

Пара функций u, v, удовлетворяющих системе (6), образует аналитическую функцию комплексного переменного $z = x + iy$

$$w = u + iv = f(z).$$

На эти функции классическими работами Коши был распространен весь аппарат дифференциального и интегрального исчислений. Этим самым этот аппарат оказался приложенным к построению и изучению конформных отображений. За истекшее столетие существования теории аналитических функций и конформных отображений получено огромное количество предложений, характеризующих различные аналитические и геометрические свойства конформных отображений. При этом оказалось, что многие из этих свойств вытекают из очень простых по формулировке геометрических принципов. Такими принципами являются так называемый принцип Шварца — Линделефа и принцип компактности.

Развитые в начале нашего столетия качественные методы теории функций действительного переменного и топологии получили значительное развитие и в теории функций комплексного переменного. С 20-х годов начали появляться исследования, направленные к созданию чисто геометрических методов в теории конформных отображений. Именно эти методы, вскрывшие сущность геометрических принципов теории конформных отображений, дали возможность распространить эту теорию на более широкие классы отображений.

7*

2. Отображения с ограниченным искажением. Первым и наиболее общим классом отображений, на который удалось распространить ряд важных свойств конформных отображений, явился класс отображений с ограниченным искажением.

Отображение (1) области D на область Δ называется *отображением с ограниченным искажением,* если характеристика отображения ограничена в D некоторой постоянной M

$$1 \leqslant p \leqslant M = \frac{1}{q}. \tag{7}$$

Класс отображений, удовлетворяющих условию (7), очевидно, инвариантен относительно конформного отображения. В силу этого достаточно ограничиться случаем, когда области D и Δ суть единичные круги с центрами в точках $z = w = 0$, причем точка $z = 0$ при отображении переходит в точку $w = 0$, т. е. $f(0) = 0$.

При этих условиях имеет место следующее основное неравенство:

$$|f(z+h) - f(z)| < K|h|^q, \tag{8}$$

где K — числовая постоянная (Грётш, Лаврентьев, Альфорс).

Сформулированное предложение характеризует равностепенную непрерывность рассматриваемого класса отображений и, следовательно, его компактность. Отсюда как следствие получается обобщение на этот класс известного предложения Каратеодори о том, что при конформном отображении областей, ограниченных жордановыми кривыми, соответствие может быть распространено на границы областей и это граничное соответствие будет взаимно непрерывным.

Опираясь на (8) и на ряд свойств конформных отображений, можно установить также следующую теорему (Лаврентьев): пусть отображение (1) удовлетворяет условию (7) и переводит единичный круг в единичный круг, причем

$$f(0) = 0, \qquad f(1) = 1;$$

при этих условиях существует функция $\mu(q)$, обращающаяся в нуль при $q = 1$, удовлетворяющая равенству

$$\lim_{q \to 1} \mu(q) = 0$$

и такая, что

$$|f(z)-z|<\mu(q).$$

При тех же условиях пусть C есть произвольная окружность в круге $|z|<1$, а γ— ее образ в круге $|w|<1$, тогда относительное отклонение γ от окружности будет меньше $\mu(q)$.

Таким образом, близость характеристики p к единице влечет за собой близость отображения к конформному. На основе этого предложения могут быть установлены сходимости различных приближенных методов построения конформных и квазиконформных отображений. Интересно отметить, что эти предложения могут быть распространены на случай отображений трехмерных областей (Лаврентьев, Крейнес).

К тому же кругу идей относятся результаты по выяснению связи между дифференцируемостью характеристик и самих функций, осуществляющих отображение. Ограничиваясь опять случаем отображения единичного круга на единичный круг, отметим следующее предложение:

Если характеристики p и t отображения (1), рассматриваемые как функции точки z, удовлетворяют условию Гёльдера с показателем ν

$$|p(z+h)-p(z)|<K|h|^{\nu}, \qquad |t(z+h)-t(z)|<K|h|^{\nu}, \qquad \nu\leqslant 1,$$

то частные производные u_x, u_y, v_x, v_y также удовлетворяют условию Гёльдера с любым показателем, меньшим ν (Лаврентьев).

Этот результат может быть усилен, если $\nu<1$ (Шабат). В этом случае частные производные будут удовлетворять условию Гёльдера с показателем ν.

3. Общая задача квазиконформных отображений. Пусть нам дана система дифференциальных уравнений с частными производными первого порядка

$$\begin{aligned} &\Phi_1\left(x, y, u, v, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}\right)=0,\\ &\Phi_2\left(x, y, u, v, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}\right)=0. \end{aligned} \tag{9}$$

Однолистное отображение

$$w=f(z), \qquad u=u(x, y), \qquad v=v(x, y) \tag{10}$$

области D на область Δ мы назовем квазиконформным отображением D на Δ, соответствующим системе (9), если функции u и v, осуществляющие это отображение, будут удовлетворять системе (9). Согласно

этому определению, квазиконформное отображение, соответствующее системе уравнений (6) Коши — Римана, будет конформным отображением. Центральной задачей теории конформных отображений, как известно, является так называемая задача Римана: даны две области, требуется построить конформное отображение одной области на другую.

В соответствии с этим, обобщенной задачей Римана теории квазиконформных отображений мы будем называть следующую задачу: *дана система уравнений* (9) *и даны две области* D *и* Δ; *требуется построить квазиконформное отображение* D *на* Δ, *соответствующее системе* (9).

Задача Римана была с исчерпывающей полнотой решена в 1913 г. Каратеодори, который показал, что если D *и* Δ односвязны и их границы содержат каждая не менее двух точек, то D можно конформно отобразить на Δ. Наряду с этим известно, что для системы

$$u_x = v_y, \qquad u_y = v_x$$

обобщенная задача Римана допускает решения для очень узкого класса областей.

Вскрыв геометрическую природу разрешимости обобщенной задачи Римана для «любых» областей, мы нашли весьма широкий класс уравнений (в некотором смысле наиболее широкий), для которого обобщенная задача Римана всегда допускает решение, причем решение единственно с точностью до трех параметров (так же как и в классической задаче Римана).

Для выявления отмеченного класса уравнений мы переформулируем задачу квазиконформных отображений в терминах характеристик отображения.

В равенствах (5), определяющих характеристики отображения, фиксируем число ν. Как было отмечено выше, частные производные u_x, u_y, v_x, v_y однозначно выражаются через характеристики

$$V,\ \alpha,\ W,\ \theta.$$

В силу этого уравнения (9) могут быть заменены двумя уравнениями, связывающими четыре характеристики и координаты соответствующих точек. Система (9) будет эквивалентна системе

$$\begin{aligned} W_\nu &= F_1^{(\nu)}(V_\nu,\ \alpha_\nu,\ x,\ y,\ u,\ v),\\ \theta_\nu &= F_2^{(\nu)}(V_\nu,\ \alpha_\nu,\ x,\ y,\ u,\ v), \end{aligned} \tag{11}$$

Систему (11) мы будем называть *уравнениями в характеристиках.* Обобщенную задачу Римана можно, таким образом, формулировать так: построить однолистное отображение области D на область Δ такое, чтобы в каждой точке D между характеристиками отображения имели место соотношения (11).

Во всем дальнейшем мы будем предполагать, что функции F_1 и F_2 однозначны, непрерывны и обладают первыми тремя непрерывными частными производными по всем аргументам.

Мы скажем, что система (11) *сильно-эллиптична,* если, каково бы ни было число ν, будут выполнены следующие условия:

1° Существует положительная постоянная k такая, что при всех значениях аргументов справедливо неравенство

$$k < \theta_\nu < \pi - k.$$

2° При любых фиксированных значениях α_ν, x, y, u, v функция F монотонно возрастает относительно ϱ_ν

$$\frac{\partial F_1}{\partial \varrho} > k > 0.$$

Заметим сейчас же, что свойство системы быть сильно эллиптичной инвариантно относительно конформного отображения; этим самым изучение квазиконформных отображений, соответствующих эллиптическим системам, сводится к изучению квазиконформных отображений круга на круг. Кроме того, из условий сильной эллиптичности непосредственно вытекает, что квазиконформные отображения, соответствующие этим системам, принадлежат к отображениям с ограниченным искажением и тем самым обладают всеми свойствами этих отображений.

Сформулируем теперь основную теорему.

Какова бы ни была сильно-эллиптическая система (11) *и каковы бы ни были две односвязные области D и Δ, ограниченные кривыми с дифференцируемыми кривизнами, всегда существует квазиконформное отображение D на Δ, соответствующее данной системе* (11). *Отображение единственно с точностью до трех параметров.*

Сформулированная основная теорема доказывается параллельно с выявлением ряда свойств квазиконформных отображений, естественно обобщающих соответствующие свойства конформных отображений,

Существенную роль при этом играет введенное нами понятие «производной системы».

4. Производные системы и свойства квазиконформных отображений. Пусть мы имеем квазиконформное отображение

$$u=u(x, y), \qquad v=v(x, y)$$

области D на область Δ, соответствующее системе (9)—(11). Рассмотрим в плоскости u, v бесконечно малый квадрат со сторонами, параллельными координатным осям, и пусть этот квадрат при нашем отображении соответствует бесконечно малому четырехугольнику q. Составим отношения длин сторон q, соответствующих сторонам квадрата, параллельным оси u и оси v. Отбрасывая при вычислениях бесконечно малые высших порядков и полагая $P=\log V$, получим

$$\begin{aligned} \frac{\partial P}{\partial v} &= a_1 \frac{\partial P}{\partial u} + a_2 \frac{\partial \alpha}{\partial u} + a_3, \\ \frac{\partial \alpha}{\partial v} &= b_1 \frac{\partial P}{\partial u} + b_2 \frac{\partial \alpha}{\partial u} + b_3, \end{aligned} \tag{12}$$

где коэффициенты a и b суть известные функции координат соответствующих точек x, y, u, v, а также характеристик P и α.

$$a_1 = \frac{\partial W}{\partial V} \operatorname{ctg} \theta - \frac{\partial \theta}{\partial V} \frac{W}{\sin^2 \theta},$$

$$a_2 = \frac{1}{V} \left[\frac{\partial W}{\partial \alpha} \operatorname{ctg} \theta - \frac{\partial \theta}{\partial \alpha} \frac{W}{\sin^2 \theta} - W \right],$$

$$a_3 = \operatorname{ctg} \theta \left[\frac{\partial W}{\partial u} \frac{1}{V} + \frac{\partial W}{\partial x} \operatorname{ctg} \alpha + \frac{\partial W}{\partial y} \sin \alpha \right] - \frac{W}{\sin^2 \theta} \left[\frac{\partial \theta}{\partial u} \frac{1}{V} + \frac{\partial W}{\partial x} \cos \alpha + \frac{\partial W}{\partial y} \sin \alpha \right],$$

$$b_1 = \frac{\partial W}{\partial V},$$

$$b_2 = \frac{1}{V} \left[\frac{\partial W}{\partial \alpha} + W \operatorname{ctg} \theta \right],$$

$$b_3 = \frac{\partial W}{\partial u} \frac{1}{V} + \frac{\partial W}{\partial x} \cos \alpha + \frac{\partial W}{\partial y} \sin \alpha.$$

Систему уравнений (12) мы будем называть *производной системой* рассматриваемого отображения.

Когда система уравнений (9) есть система уравнений Коши — Римана, производная система (12) переходит также в систему уравнений Коши—Римана. В силу этого при конформном отображении некоторой области D отображение D, осуществляемое парой характеристик

$$P = P(x, y) \text{ и } \alpha = \alpha(x, y),$$

будет также конформным.

Этот основоположный факт теории аналитических функций допускает замечательное распространение на любые квазиконформные отображения. Имеет место следующая теорема:

Если система (11) *сильно-эллиптична, то система* (12) *эллиптична в обычном смысле*: $4a_2b_1 > (b_2 - a_1)^2$, *а ее однородная часть*

$$\frac{\partial P}{\partial v} = a_1 \frac{\partial P}{\partial u} + a_2 \frac{\partial \alpha}{\partial u},$$

$$\frac{\partial \alpha}{\partial v} = b_1 \frac{\partial P}{\partial u} + b_2 \frac{\partial \alpha}{\partial u}$$

будет сильно-эллиптичной.

Из условия эллиптичности вытекает, что систему (12) можно всегда разрешить относительно $\frac{\partial P}{\partial u}$ и $\frac{\partial P}{\partial v}$

$$\begin{aligned} \frac{\partial P}{\partial u} &= A_1 \frac{\partial \alpha}{\partial u} + B_1 \frac{\partial \alpha}{\partial v} + C_1, \\ \frac{\partial P}{\partial v} &= A_2 \frac{\partial \alpha}{\partial u} + B_2 \frac{\partial \alpha}{\partial v} + C_2. \end{aligned} \tag{12'}$$

Отмечу сейчас же, что для случая рассматриваемых ниже уравнений (16) газовой динамики производная система переходит в систему уравнений Чаплыгина, полученную им аналитическим путем и легшую в основу многих фундаментальных работ по газовой динамике.

Опираясь на эту теорему и на некоторые свойства квазиконформных отображений, соответствующих линейным системам, можно установить ряд теорем, касающихся связи между дифференцируемостью функций

F_1 и F_2 и дифференцируемостью функций u, v. Приведу одну из этих теорем.

Если система (11) *сильно-эллиптична и если функции* F_1 *и* F_2 *обладают частными производными, удовлетворяющими условию Гёльдера, то при квазиконформном отображении функции* P *и* α *обладают частными производными, удовлетворяющими также условию Гёльдера.*

Теорема о производных системах лежит также в основе обобщения на квазиконформные отображения принципа Шварца — Линделёфа.

Мы сформулируем этот принцип в терминах, наиболее удобных для приложений к задаче газовой динамики; кроме того, мы ограничимся случаем, когда функции F_1 и F_2 не зависят явно от координат соответствующих точек

$$W=F_1(V,\ \alpha),\qquad \theta=F_2(V,\ \alpha). \tag{13}$$

Пусть в плоскости xOy кривые Γ_0 и Γ заданы соответственно уравнениями

$$y=y_0(x),\qquad y=Y(x),\qquad y_0(x)<Y(x).$$

Область плоскости xOy, заключенную между кривыми Γ_0 и Γ, обозначим через $D(\Gamma_0,\ \Gamma)$; кроме того, обозначим через Δ полосу, заключенную между прямыми $v=0$ и $v=h$. Функции $y_0(x)$ и $Y(x)$ мы будем предполагать определенными и дважды дифференцируемыми при всех значениях x, $|x|<\infty$.

Согласно основной теореме, существует квазиконформное отображение области D на Δ

$$w=f(z,\ \Gamma_0,\ \Gamma), \tag{14}$$

соответствующее системе (13). Свободными параметрами мы распорядимся так, чтобы точки $\pm\infty$ области D переходили в точки $\pm\infty$ области Δ

$$f(\pm\infty,\ \Gamma_0,\ \Gamma)=\pm\infty.$$

Линию области D, переходящую при отображении (14) в прямую $v=\text{const}$, мы будем называть линией тока и будем обозначать через y_v

$$y_v=y(x,\ v).$$

При этих обозначениях обобщенный принцип Шварца может быть сформулирован следующим образом.

Пусть $D(\Gamma_0, \overline{\Gamma})$ есть область, ограниченная линиями Γ_0 и $\overline{\Gamma}$, уравнения которых

$$y = y_0(x), \quad y = \overline{y}(x)$$

и

$$w = f(z, \Gamma_0, \overline{\Gamma}), \quad f(\pm\infty, \Gamma_0, \overline{\Gamma}) = \pm\infty,$$

функция, осуществляющая квазиконформное отображение $\overline{D}(\Gamma_0, \overline{\Gamma})$ на полосу B, соответствующее системе (13). Если $\overline{y}(x) \geqslant y(x)$ и система (13) сильно-эллиптична, то при любом v

$$\overline{y}(x, v) \geqslant y(x, v).$$

Знак равенства в последнем неравенстве, при $v < h$, может быть только при условии $\overline{y}(x) = y(x)$.

Таким образом, все свойства конформных отображений, опирающихся на принцип Шварца, автоматически распространяются на квазиконформные отображения, соответствующие сильно-эллиптическим системам (13). В частности этот принцип позволяет на изучаемый класс распространить теорию граничных производных. Отметим один результат, допускающий для случая уравнений газовой динамики наглядную механическую иллюстрацию.

Пусть в точке x_0 производная $y'(x)$ достигает абсолютного максимума и пусть при всех значениях x

$$y_0'(x) < y'(x_0);$$

при этих условиях производная характеристики V в точке x_0, $y(x_0)$ по дуге Γ в направлении возрастающих x всегда положительна.

Принцип Шварца может быть также сформулирован в несколько суженной форме для наиболее общих квазиконформных отображений, соответствующих сильно-эллиптическим системам.

Отметим в заключение одну формулу, которая может быть положена в основу фактического конструирования квазиконформного отображения. Мы ограничимся опять случаем системы (13).

Сохраняя обозначения, принятые выше, допустим, что линии Γ_0 и Γ близки, а число h мало

$$h < Y(x) - y_0(x) < Ch,$$
$$|Y'(x) - y_0'(x)| < ch,$$
$$|Y''(x) - y_0''(x)| < ch,$$
$$|Y'''(x)| + |y'''(x)| < c,$$

где c и C—постоянные, не зависящие от h. Пусть теперь z—произвольная точка Γ; обозначим через $n = n(z)$ длину отрезка нормали Γ, выходящего из z и заключенного между Γ и Γ_0, через $K(z)$ обозначим кривизну Γ в той же точке

$$K = \frac{Y''}{(1 + Y'^2)^{3/2}}$$

и через $\alpha = \alpha(z)$—угол наклона касательной к Γ в точке z: $\alpha = \operatorname{arc\,tg} Y'(x)$, а через α_0 — угол между касательными к Γ_0 и Γ, проведенными через концы отрезка n.

Отобразим область $D(\Gamma_0, \Gamma)$ на полосу Δ

$$w = f(z, \Gamma_0, \Gamma), \qquad f(\pm\infty, \Gamma_0, \Gamma) = \pm\infty$$

и обозначим через $V = V(z)$ значение характеристики V в точке z, а через V_0 ее приближенное значение, определяемое уравнением

$$W_0 = \frac{n}{h} = F\left(V_0, \frac{\alpha + \alpha_0}{2}\right).$$

При этих обозначениях получим в точке z следующую формулу

$$\log V(z) = \log V_0 + \frac{1}{2}\left[A_2 V_0 h - \left(A_1 - B_2 + B_1 \operatorname{ctg}\theta \frac{W_0}{V_0}\right)\operatorname{ctg}\theta_n\right]K + $$
$$+ \frac{1}{2}\left[B_2 - B_1 \operatorname{ctg}\theta \frac{W_0}{V_0}\right](\alpha - \alpha_0) + Rh^2, \qquad (15)$$

где R остается ограниченным при h стремящемся к нулю, и где A и B — коэффициенты производной системы (12), причем в выражениях этих коэффициентов, а также угла $\theta = F_2(\varphi, \alpha)$, вместо φ и α, следует подставить φ_0 и $\alpha(z)$.

Если в формуле (15) отбросить последний член, то мы получим приближенную формулу для вычисления V с точностью до малых первого порядка сравнительно с шириной полос.

При квазиконформном отображении полос конечной ширины на прямолинейные полосы можно, опираясь на формулу (15), построить метод для приближенного определения линий тока $y=y(x, v)$. Функции $y(x, v_i)$, где $i=1, 2, \ldots, n$, будут определяться как интегральные кривые системы n обыкновенных дифференциальных уравнений второго порядка.

5. Квазиконформные отображения и задачи газовой динамики. Основные задачи плоского дозвукового установившегося движения идеального газа можно рассматривать как частные случаи общей задачи квазиконформных отображений.

Рассмотрим, для определенности, задачу построения газового потока, двигающегося между двумя цилиндрическими стенками. Возьмем плоскость xOy, $z=x+iy$, перпендикулярную к образующим стенок, и пусть Γ_0 и Γ, уравнения которых

$$y=y_0(x), \qquad y=Y(x),$$

суть линии пересечения стенок с этой плоскостью. Для решения задачи достаточно найти потенциал скоростей $u(x, y)$ течения, который, как известно, вместе с функцией тока $v(x, y)$ должен удовлетворить следующей системе уравнений с частными производными:

$$\begin{aligned} \frac{\partial u}{\partial x} &= p\frac{\partial v}{\partial y}, \\ \frac{\partial u}{\partial y} &= -p\frac{\partial v}{\partial x}, \end{aligned} \tag{16}$$

где коэффициент p есть заданная функция скорости потока

$$p=p(t)=\left(1-\frac{t^{-2}}{2\alpha}\right)^{-\beta},$$

$$t^{-1}=\sqrt{\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2},$$

а α и β — положительные величины, просто выражающиеся через газовые постоянные.

По условию задачи стенки должны состоять из линий тока, следовательно, нужно искать решение системы (16), при котором на линиях Γ_0 и Γ функция v должна сохранять постоянное значение. Вдоль Γ_0 можно считать $v=0$; пусть вдоль Γ будет $v=h$. Число h должно быть задано заранее, ибо оно определяет расход, т. е. количество газа, протекающего через сечение потока, отнесенное к единице времени и высоте. Таким образом, искомое решение системы (16)

$$\left.\begin{aligned} u&=u\ (x,\ u) \\ v&=v\ (x,\ u) \end{aligned}\right\}\ w=f\ (z)$$

должно давать отображение области $D\ (\Gamma_0,\ \Gamma)$ на полосу Δ: $0<v<h$. Сформулированная задача газовой динамики эквивалентна задаче квазиконформного отображения D на Δ, соответствующего системе (16). Наша система примет особо простой вид, если ее записать в характеристиках

$$W=Vp\,(V)=F\ (V),\qquad \theta=\frac{\pi}{2}\ . \tag{17}$$

Заметим, что в рассматриваемом случае характеристика отображения V есть величина, обратная скорости потока.

Анализируя выражение для F, нетрудно показать, что знак производной F зависит от V: при V^{-1} меньше скорости звука $F'\ (V)>0$, при V^{-1} больше скорости звука $F'\ (V)<0$.

Если в газовом потоке скорости всюду меньше звуковой, то система (17) сильно-эллиптична и к ней могут быть применены все общие теоремы теории квазиконформных отображений, которые в этом случае приобретают механическую наглядность. Начнем с теоремы существования.

Допустим, что линии Γ_0 и Γ обладают ограниченной кривизной, а расстояние между Γ_0 и Γ ограничено снизу. При этих условиях всегда найдется такое h_0, что для всех расходов $h\leqslant h_0$ можно построить газовый поток в полосе $D\ (\Gamma_0,\ \Gamma)$, причем при $h=h_0$ в потоке найдется точка, где скорость потока будет равна скорости звука.

Замечу, что в рассматриваемой задаче точного решения аналитические методы разложений по параметру (Келдыш и Франкль) дают

возможность строить поток лишь при достаточно малых скоростях — достаточно малых h.

За последнее десятилетие исключительное развитие получили различные приближенные методы для построения газовых потоков (методы Христиановича). Для уточнения этих решений, а также для решения более сложных газодинамических задач, связанных с поисками границ потоков, удовлетворяющих специальным условиям, может быть использован обобщенный принцип Шварца, который в терминах газовой динамики может быть сформулирован следующим образом.

Пусть мы имеем в полосе D (Γ_0, Γ) газовый дозвуковой поток с расходом h; если мы, сохраняя h, проварьируем часть стенки Γ_0 вовне потока — вниз, то все линии тока опустятся, скорости потока на Γ всюду уменьшатся, на недеформированной части Γ_0 скорости увеличатся, а в месте наибольшего искажения Γ_0 скорости уменьшатся.

Из более сложных краевых задач уравнений газовой динамики отмечу решенную нами задачу на обтекание со срывом струй: пусть стенки Γ_0 и Γ симметричны относительно оси

$$-y_0(x) = +Y(x) > K > 0$$

и пусть в области D (Γ_0, Γ) дана дуга γ

$$y = \pm\varphi(x), \qquad 0 \leqslant x \leqslant a,$$

расположенная строго внутри D (Γ_0, Γ) и также симметричная относительно оси Ox. Если кривизна Γ_0, Γ и γ ограничена, то для всех дозвуковых режимов можно в области D (Γ_0, Γ) построить газовый поток, обтекающий γ со срывом струй в концах γ.

6. Линейные системы и их приложения. Несколько в ином направлении за последнее десятилетие получила развитие задача квазиконформных отображений, соответствующих линейным системам

$$\frac{\partial u}{\partial x} = a_1 \frac{\partial v}{\partial x} + b_1 \frac{\partial v}{\partial y}, \qquad (18)$$
$$\frac{\partial u}{\partial y} = a_2 \frac{\partial v}{\partial x} + b_2 \frac{\partial v}{\partial y},$$

где a и b — заданные функции x и y. В этом случае, предполагая a и b дифференцируемыми, можно из системы уравнений (18) исключить функцию u; для v мы получим линейное уравнение с частными производными второго порядка; задача квазиконформного отображения редуцируется к некоторой краевой задаче для одного линейного уравнения, для решения и изучения которого существуют достаточно мощные аналитические методы. Именно в этом направлении, с использованием аппарата интегральных уравнений, около 30 лет тому назад было дано решение геометрической задачи о конформном отображении заданного куска поверхности на плоский круг (Лихтенштейн). Для применимости метода пришлось потребовать, чтобы в каждой точке отображений поверхности существовала касательная плоскость и чтобы направляющие углы нормали к этой плоскости при перемещении точки касания были не только непрерывны, но удовлетворяли бы дополнительно условию Гёльдера.

Применяя к задаче конформного отображения поверхностей (Лаврентьев) геометрические принципы квазиконформных отображений, удалось получить предельно простое решение этой задачи в наиболее общих условиях — без дополнительного условия Гёльдера. При этом обнаружилось, что построенное решение в общем случае удовлетворяет условиям конформности, но не всюду дифференцируемо.

Эти результаты были впоследствии распространены на квазиконформные отображения, соответствующие любым эллиптическим системам (18), при единственном условии непрерывности коэффициентов a и b (Шабат). Основную теорему существования удалось также распространить на случай, когда a и b непрерывным образом зависят не только от независимых переменных x, y, но также и от функций u, v (Шапиро).

7. Заключение. В приведенном очерке я по возможности подробно остановился на наиболее принципиальных моментах теории и лишь слегка коснулся ее некоторых приложений. В частности я не упомянул о ряде существенных приложений линейных задач к теории функций комплексного переменного.

Отмечу в заключение два больших направления, в которых, как мне кажется, должны в первую очередь развиться изложенные выше результаты.

Первое направление должно относиться к распространению теории на отображения трехмерных областей и в первую очередь на отображения, связанные с трехмерными течениями идеальной жидкости и газа. Имеющиеся здесь результаты хотя и носят пока частный характер, но показывают принципиальную возможность такого распространения.

Второе направление должно относиться к задачам квазиконформных отображений, соответствующим уравнениям смешанного типа. Замечательные результаты, полученные в этом направлении за последние годы для уравнений газовой динамики С. А. Христиановичем, а также Ф. И. Франклем, показывают на наличие здесь многих глубоких закономерностей, которые должны распространиться на весьма широкие классы уравнений смешанного типа.

Академик
С. Н. Бернштейн

О РОЛИ НЕРАВЕНСТВ И ЭКСТРЕМАЛЬНЫХ ПРОБЛЕМ В МАТЕМАТИЧЕСКОМ АНАЛИЗЕ

§ 1. Одной из проблем, которые привели к созданию в XVII веке анализа бесконечно малых, была задача нахождения экстремальных значений, т. е. максимумов и минимумов, характеризующих ход изменения заданной функции в той или иной области. В XVIII веке, в связи с новыми проблемами геометрии и механики, возникло так называемое вариационное исчисление, имевшее целью решение экстремальных задач более общего характера, образцом которых может служить отыскание геодезических линий данной поверхности, т. е. нахождение кратчайшего из путей, расположенных на этой поверхности, соединяющих две фиксированные на ней точки.

Каждый студент знает и неоднократно применял для отыскания экстремумов данной функции известный метод дифференциального исчисления, заключающийся в приравнивании нулю ее производной. Однако даже при исследовании сравнительно простых функций уравнения, к которым приводит этот метод, часто не допускают формального решения при помощи конечного числа известных алгебраических операций, и нередко умелое использование знака производной быстрее и проще приводит к цели. Кроме того, единственными общими эффективными приемами решения каких бы то ни было численных уравнений являются соответствующие методы последовательных приближений.

Еще обычнее случаи, когда решения дифференциальных уравнений, и в частности тех, к которым приводит классический метод вариационного исчисления, не могут быть формально выражены при помощи конечного числа ранее встречавшихся математических функций. Таким образом, самый вопрос о существовании функции, удовлетворяющей всем требованиям поставленной задачи, остается открытым.

Ввиду этого уже Ньютон, Эйлер и другие классики широко пользовались бесконечными процессами, например суммами бесконечных рядов, члены которых последовательно составляются при помощи конечного числа арифметических действий, однозначно определяемых условиями решаемой задачи.

Но этот замечательный своей общностью «трансцендентный» метод представления функций,— дававший в некотором обобщенном смысле формальное решение самых разнообразных задач анализа и его приложений, и после соответствующей переработки, связанной с систематическим применением неравенств, занявший центральное место в современной математике,— обладал, однако, противоположным опасным недостатком — чрезмерной универсальностью: искомой функции соответствует бесконечный ряд, формально удовлетворяющий требованиям задачи даже и тогда, когда этим требованиям не может удовлетворить никакая функция. Поэтому вопрос о существовании решения и его эффективном вычислении все же оставался открытым. Не трудно проверить, например, применяя безупречный конечный формальный метод, что решением уравнения Лапласа

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0,$$

удовлетворяющим при $y=0$ начальным условиям (Коши—Ковалевской)

$$f(x, 0) = \sum_{k=0}^{n} A_k \cos kx + B_k \sin kx \text{ и } \frac{\partial f(x, 0)}{\partial y} = 0,$$

является периодическая по x функция

$$f(x, y) = \sum_{k=0}^{n} \frac{e^{ky} + e^{-ky}}{2} (A_k \cos kx + B_k \sin kx),$$

8*

каковы бы ни были заданные $2n+1$ коэффициентов A_k, B_k. Применение же упомянутого выше бесконечного формального метода, соответствующее случаю, когда начальная функция $f(x, 0)$ выражается суммой членов *бесконечного* сходящегося ряда

$$f(x,\ 0) = \sum_0^\infty A_k \cos kx + B_k \sin kx$$

приводит к аналогичному ряду

$$f(x,\ y) = \sum_{k=0}^\infty \frac{e^{ky}+e^{-ky}}{2}(A_k \cos kx + B_k \sin kx).$$

Между тем известно, что уравнение Лапласа не может иметь решения, если $\left(\text{при } \frac{\partial f}{\partial y}(x,\ 0)=0\right)$ начальная функция $f(x,\ 0)$ *неаналитическая*[1] $\Big($например, если $f(x,\ 0) = |\cos x| = \frac{2}{\pi}\left\{1 + \frac{2}{3}\cos 2x - \frac{2}{15}\cos 4x + \ldots \pm \pm \frac{2}{4k^2-1}\cos 2kx + \ldots\right\}\Big)$: легко видеть, что в последнем случае построенный выше ряд, определяющий функцию $f(x,\ y)$, оказывается расходящимся при любом $y \gtrless 0$.

Вообще для того, чтобы ряд, формально решающий задачу, мог быть непосредственно использован для сколь угодно точного приближенного вычисления значений искомой функции при заданных значениях переменных, необходимо, чтобы полученный ряд был сходящимся в соответствующей области. Поэтому после величайшего расцвета математического конструктивного творчества в XVIII веке одним из главных устремлений математиков последующих поколений должно было стать уточнение условий законности указанных выше «трансцендентных» построений, которое в первую очередь требовало установления условий сходимости бесконечных процессов (рядов, произведений, ин-

[1] Функция $f(x)$ называется аналитической в промежутке $a < x < b$, если в каждой точке x_0 этого промежутка функция $f(x)$ может быть разложена в ряд Тэйлора по степеням $(x-x_0)$, сходящийся при каком-нибудь значении $x \neq x_0$ (функция $|\cos x|$ *не* может быть разложена в степенной ряд при $x=\frac{\pi}{2}$).

тегралов), применявшихся при решении различных задач анализа и его приложений. Отсюда по необходимости возникают экстремальные проблемы нового типа, связанные с оценкой погрешности, получаемой при выполнении лишь определенной части данного бесконечного процесса (например, при вычислении суммы конечного числа членов ряда).

§ 2. Для иллюстрации характера этих проблем напомню, как формальное решение систем обыкновенных дифференциальных уравнений посредством бесконечных степенных рядов привело Коши к постановке и решению экстремальной проблемы, на котором основано его классическое доказательство существования и вычисление функций, удовлетворяющих данной системе уравнений при соответствующих начальных данных. Рассмотрения одного уравнения первого порядка достаточно, чтобы понять, в чем заключается сущность экстремального метода доказательства Коши [1].

Предполагая, что уравнение

$$\frac{dy}{dx}=f(x, y),$$

где $f(x, y)=\sum B_{ik}x^i y^k$ выражается степенным рядом, сходящимся при $|x|\leqslant r$, $|y|\leqslant R$ (т. е. существует такая постоянная M, что $|B_{ik}|\leqslant \leqslant\frac{M}{r^i R^k}$), имеет решение

$$y=\varphi(x)=\sum_1^\infty a_k x^k,$$

которое обращается в нуль при $x=0$, для вычисления и установления сходимости $\varphi(x)$ нужно определить верхние грани последовательных коэффициентов a_k (т. е. последовательных производных $\varphi(x)$ при $x=0$). Коши замечает, что $|a_k|$ достигают наибольшего значения, когда все B_{ik} положительны, и таким образом $|a_k|\leqslant A_k$, если $\Phi(x)=\sum_1^\infty A_k x^k$ есть

[1] Этот метод, который Коши называл «Calcul des limites» (исчисление границ), теперь обычно называют способом майорант.

решение (майоризирующего) дифференциального уравнения

$$\frac{dy}{dx} = \frac{M}{\left(1 - \frac{x}{r}\right)\left(1 - \frac{y}{R}\right)},$$

которое легко получается в конечном виде посредством разделения переменных.

Метод майорант был затем применен Софьей Ковалевской к доказательству существования решения уравнений в частных производных при так называемых условиях Коши, в предположении, конечно, что все начальные функции, как и само уравнение, аналитические.

Независимо от формального аппарата, который впоследствии применялся при решении и исследовании различного рода дифференциальных или функциональных уравнений, все эффективные общие доказательства существования решений, удовлетворяющих некоторым определенным условиям, имеют в своей основе экстремальный метод, сходный по идее с методом Коши. Например, даже наиболее глубоко отличный по форме метод последовательных приближений Пикара, не предполагающий аналитичности функции $f(x, y)$, основан на определении верхней грани $|y|$ и $\left|\frac{dy}{dx}\right|$, когда задана верхняя грань $|f(x, y)|$ при $|x| \leqslant r$, $|y| \leqslant R$.

Областью сходимости степенного ряда, являющегося простейшим бесконечным функциональным процессом, всегда служит круг, находящийся в комплексной плоскости. Таким образом, благодаря технике способа майорант, который нисколько не связан с вещественностью начальных данных, систематическое введение комплексной переменной для полного изучения и построения функций, удовлетворяющих аналитическим дифференциальным уравнениям, представлялось в прошлом столетии не только полезным, но и столь же необходимым, как и в теории алгебраических уравнений или при построении некоторых важных новых функций, например эллиптических, открытых еще в конце XVIII века.

В самом деле, аналитическая теория дифференциальных уравнений достигла замечательных успехов, открыв новые методы построения во всей комплексной плоскости функций комплексной переменной, удовлетворяющих некоторым важным классам дифференциальных уравнений, в частности линейным. Однако дифференциальные уравнения,

исследование которых во всей комплексной плоскости осуществимо при современном состоянии науки, составляют лишь незначительную часть всего многообразия дифференциальных уравнений, встречающихся в анализе и его приложениях.

Ввиду этого в конце прошлого столетия возникло новое вещественное направление, основоположниками которого были Пуанкаре и Ляпунов, имеющее целью изучение и построение всего семейства *вещественных* кривых, удовлетворяющих данной системе обыкновенных дифференциальных уравнений. По самой постановке вопроса центральное место в этой теории занимают неравенства; при этом основным математическим аппаратом исследования остаются аналитические функции вещественной переменной, а именно: существенную роль играет метод, родственный теории пертурбаций небесной механики, заключающийся в введении вспомогательного параметра, по степеням которого раскладываются искомые решения. Как увидим далее, этот параметрический метод был впоследствии с успехом применен и в других областях анализа.

§ **3.** Если теория функций комплексной переменной заняла во второй половине прошлого столетия господствующее место в развитии математики, то этим она обязана в не меньшей степени, чем важным проблемам анализа, естественно вовлеченным в ее орбиту, исключительной стройности и мощи своих методов.

Как известно, в основе теории функций комплексной переменной лежит классическая интегральная формула

$$f(x)=\frac{1}{2\pi i}\int\limits_C \frac{f(z)\,dz}{z-x},$$

открытая тем же великим французским математиком Коши, которая выражает регулярную внутри данной области функцию $f(x)$ посредством ее значений на границе C этой области.

Вследствие своей замечательной гибкости (вытекающей из широкой свободы выбора контура интегрирования C) формула Коши стала неиссякаемым источником формальных трансцендентных тождеств и представлений функций различных классов в виде рядов, сходящихся во всех точках, где рассматриваемая функция регулярна. Благодаря этому, например, степенной ряд, практически дающий решение дифференциаль-

ного уравнения обычно лишь вблизи начальной точки, может быть преобразован в сходящийся ряд многочленов типа Миттаг-Лефлера, осуществляющий аналитическое продолжение искомой функции, т. е. позволяющий ее вычислить в любой точке области регулярности (в частности, на всем вещественном отрезке, где функция аналитична).

Но, несмотря на то, что все свойства функции потенциально заложены в коэффициентах ее разложения в строку Тэйлора в какой-нибудь данной точке регулярности, тем не менее невозможно указать единого принципиально общего метода для определения области регулярности любой заданной таким образом функции и характеристики всех ее особенностей, являющихся в некотором смысле инвариантами бесконечной группы преобразований, имеющей элементами строки Тэйлора рассматриваемой функции во всех точках ее регулярности. Формальная теория групп в ее современном абстрактном аспекте мало содействовала в последние годы развитию теории функций комплексной переменной и, в частности, продолжению крупных успехов, достигнутых в конце прошлого столетия Пуанкаре и другими в изучении многозначных функций и их конструктивной униформизации. Наиболее плодотворное развитие за последние десятилетия получили те разделы теории функций комплексной переменной, в которых, как в теории целых и мероморфных функций, руководящее место заняли неравенства и связанные с ними различные экстремальные проблемы.

Напомню здесь лишь фундаментальные неравенства, свойственные всем регулярным функциям комплексной переменной, среди которых элементарные многочлены $(Ax+B)^k$ выделяются соответствующим характеризующим их экстремальным свойством. *Если функция* $f(x)$ *регулярна и ограничена в круге* C *с центром* a *и радиусом* R $(|f(x)|\leqslant M)$, *то все ее производные* $f^{(k)}(x)$ $(k=1, 2, \ldots)$ *в центре* a *удовлетворяют неравенствам* $\left|f^{(k)}(a)\right|\leqslant\frac{Mk!}{R^k}$, *причем (для каждого данного* k*) знак равенства осуществляется, когда* $f(x)=M\frac{(x-a)^k}{R^k}$.

Эта теорема, которая легко выводится из указанной выше интегральной формулы Коши, также есть следствие того, что вещественная и мнимая части функции комплексной переменной удовлетворяют уравнению (эллиптического типа) Лапласа.

§ 4. Замечательно, что аналогичным фундаментальным неравенствам удовлетворяют решения всех аналитических дифференциальных уравнений второго порядка в частных производных эллиптического типа и, в частности, уравнения (Эйлера — Лагранжа), к которым приводит упомянутая в начале общая проблема вариационного исчисления. Однако в общем случае (нелинейных уравнений) соответствующим неравенствам удовлетворяют не модули рассматриваемых функций, а некоторые их технические видоизменения (нормы или нормализованные модули этих функций), специально приспособленные к исследованию аналитических функций в вещественной области, на построении которых нет надобности здесь останавливаться. Благодаря полученным таким образом неравенствам, в начале нынешнего столетия удалось установить математический факт крупнейшего принципиального значения, предугаданный Гильбертом, что всякая функция, являющаяся решением экстремальной задачи вариационного исчисления, должна быть аналитической[1]. Иными словами, аналитические функции вещественных переменных представляются теперь как особый класс функций, характеризуемый присущими ему вещественными свойствами, которому, в частности, принадлежат все функции, возникающие из проблем вариационного исчисления, не умещающиеся, однако, в прокрустово ложе общей классической теории функций комплексной переменной, так как область их регулярности обычно ограничивается вещественными и достаточно близкими к ним комплексными значениями переменных.

Неравенства указанного типа позволили также произвести полное исследование условий существования решения первой краевой задачи (Дирихле) для любого эллиптического уравнения и дать общий метод для эффективного его вычисления. Тот же метод с теми или иными техническими изменениями может быть применен при решении любых других краевых задач для уравнений эллиптического типа.

Сущность этого (параметрического) метода заключается в следующем. В данное уравнение

$$F\left(\frac{\partial^2 z}{\partial x^2},\ \frac{\partial^2 z}{\partial x \partial y},\ \frac{\partial^2 z}{\partial y^2},\ \frac{\partial z}{\partial x},\ \frac{\partial z}{\partial y},\ z,\ x,\ y\right)=0, \tag{1}$$

[1] Каковы бы ни были краевые задания (при условии, что подинтегральная функция в обращаемом в минимум интеграле аналитическая).

где F — аналитическая функция относительно всех входящих в нее переменных, вводится параметр α так, чтобы при $\alpha=1$ оно обращалось в (1), а при $\alpha=0$

$$F\left(\frac{\partial^2 z}{\partial x^2}, \ldots, x, y, \alpha\right)=0 \tag{2}$$

обращалось бы в такое уравнение (например, Лапласа), для которого решение $z(x, y, 0)$ рассматриваемой при тех же краевых данных задачи известно. В таком случае при некоторых весьма общих предположениях относительно гладкости заданных краевых функций устанавливается при помощи упомянутых выше фундаментальных неравенств, что для достаточно малых α ($|a|\leqslant r$) решение $z(x, y, \alpha)$ уравнения (2), удовлетворяющее тем же краевым заданиям, также существует и является аналитической функцией параметра α, выражаемой степенным рядом (Тэйлора)

$$z(x, y, \alpha)=\sum_0^\infty \frac{\partial^k z(x, y, 0)}{\partial \alpha^k} \frac{\alpha^k}{k!}, \tag{3}$$

коэффициенты которого вычисляются последовательно без труда, как решения соответствующих уравнений Пуассона. Кроме того, благодаря тем же неравенствам функция $z(x, y, \alpha)$ при дальнейшем возрастании α, вплоть до значения $\alpha=1$, продолжает удовлетворять уравнению (2), оставаясь *аналитической по α, если из краевого условия вытекает равномерная ограниченность $z(x, y, \alpha)$ и его частных производных первых двух порядков по x и y*. Следовательно, при этом условии, согласно сказанному ранее, ряд Миттаг-Лефлера, преобразованный из степенного ряда (3), сходится на отрезке (0, 1) регулярности относительно α функции $z(x, y, \alpha)$ и при $\alpha=1$ дает искомое решение $z(x, y)=z(x, y, 1)$ уравнения (1). Последнее условие (возможности ограничения z и его производных) является по существу также и необходимым для существования решения рассматриваемой краевой задачи.

Применение параметрического метода, который в указанной выше форме приводит решение краевой задачи к аналитическому продолжению, обусловлено аналитической природой уравнения (1). Но благодаря теореме Вейерштрасса о представимости всякой непрерывной функции в виде равномерно сходящегося ряда многочленов высказан-

ное условие существования решения распространяется также и на неаналитические уравнения: «таким образом, ограничение последовательных частных производных составляет истинную сущность задачи Дирихле».[1] Приведенная здесь краткая цитата из введения к моей магистерской диссертации имеет широкое принципиальное значение. Например, невозможность существования решения $z(x, y)$ уравнения Лапласа, удовлетворяющего рассмотренным вначале (§ 1) условиям Коши $\left(\text{когда } \frac{\partial z(x, 0)}{\partial y}=0,\ z(x, 0) - \textit{неаналитическая} \text{ функция}\right)$, связана с невозможностью указать верхнюю грань $|z(x, y)|$ ни при каком $y \gtrless 0$, когда заданная функция

$$z(x, 0)=\sum_{0}^{n} A_k \cos kx + B_k \sin kx$$

выражается (как и ее решение, указанное в § 1) *конечной* тригонометрической суммой, с произвольным n, если известны только верхние грани любого данного конечного числа модулей ее последовательных частных производных $\left(\frac{\partial^i z(x,0)}{\partial x^i}\right)$. Распространение высказанной идеи на другие типы уравнений осуществлено пока лишь в малой степени и может служить основой программы будущих исследований.

§ 5. Наиболее совершенным математическим аппаратом, который достался нашему поколению в наследие от прошлого столетия, была теория функций комплексной переменной. Поэтому, подводя 40 лет тому назад итоги развития анализа, я имел основания сказать, учитывая намечавшиеся тогда новые течения, что «теория аналитических функций впервые открывает. . . определенный, хотя и не всегда кратчайший путь к разрешению всякой задачи анализа — аналитическое продолжение».[2]

Если современная математика находит более короткие пути к решению ряда вопросов, прямо или косвенно связанных с естествознанием и техникой, то этим она обязана в большой степени новому направле-

[1] С. Н. Бернштейн. Исследование и интегрирование дифференциальных уравнений с частными производными второго порядка эллиптического типа. Сообщ. Харьковского матем. об-ва, 1908.

[2] Конец введения к той же работе.

нию, объединяющему теорию всех математических функций на базе свойств их приближения посредством тех или иных более простых функций, идейным основоположником которого является в первую очередь П. Л. Чебышев. Именно он обратил внимание на огромную важность и для практики и для дальнейшего развития анализа решения разного рода экстремальных задач, возникающих при математической трактовке вопроса о кратчайшем или наилучшем в некотором определенном смысле, «наивыгоднейшем», способе достижения поставленной цели.

Одной из важнейших математических проблем, которая стала таким образом перед Чебышевым, была замена многочлена степени n, состоящего из первых $(n+1)$ членов степенного ряда по степеням $(x-c)$, хорошо приближающего данную функцию $f(x)$ лишь *вблизи* точки c, таким многочленом $P_n(x)$ той же степени n, который позволил бы вычислить функцию $f(x)$ с наименьшей погрешностью в данном промежутке (a, b) любой длины. Решение этой задачи определения многочлена $P_n(x)$ степени n, наименее уклоняющегося от функции $f(x)$ в данном промежутке (a, b), не может быть получено в конечном виде, за исключением очень редких случаев [например, когда $f(x)= \frac{A}{x-d}$, где d — данная точка *вне* отрезка (a, b)]. Среди известных в настоящее время сходящихся процессов последовательных приближений, разрешающих эту экстремальную задачу для любой данной функции, следует отметить прием, представляющий соответствующее приспособление указанного выше параметрического метода, в основу которого здесь положена фундаментальная теорема Чебышева о признаке, характеризующем многочлен $P_n(x)$, наименее уклоняющийся от данной функции $f(x)$ в рассматриваемом промежутке (a, b).

Однако ввиду сравнительной сложности всех общих методов вычисления многочленов, наименее уклоняющихся от данной функции, особое значение приобретает установление неравенств, которым удовлетворяет наименьшая возможная погрешность на (a, b), т. е. наилучшее приближение $E_n(f(x); a, b)$ функции $f(x)$, зависящее существенным образом от природы дифференциальных свойств рассматриваемой функции. В самом деле, благодаря этим неравенствам мы можем сказать заранее, возможен ли многочлен степени n, дающий настолько лучшее

приближение рассматриваемой функции, чем то, которое осуществляет некоторый уже вычисленный нами многочлен той же степени, чтобы дальнейшее уточнение вычислений было целесообразным.

Легко показать, например, что

$$E_n(f(x);\ a,\ b) \leqslant \frac{2M}{(n+1)!}\left(\frac{b-a}{4}\right)^{n+1}, \tag{4}$$

если известно, что

$$|f^{(n+1)}(x)| \leqslant M \text{ при } a \leqslant x \leqslant b; \tag{5}$$

напротив,

$$E_n(f(x);\ a,\ b) \geqslant \frac{2M}{(n+1)!}\left(\frac{b-a}{4}\right)^{n+1},$$

если имеет место противоположное неравенство $f^{(n+1)}x \geqslant M$. Отсюда, между прочим, видно, что в случае непрерывности $f^{(n+1)}(x)$ знак равенства в (4) фактически осуществляется только одновременно со знаком равенства в (5) при всех x $(a \leqslant x \leqslant b)$, т. е. когда $f(x)$ есть многочлен степени[1] $n+1$.

Установление подобного рода неравенств потребовало прежде всего глубокого изучения экстремальных свойств алгебраических многочленов в вещественной области, так как проблема наилучшего приближения произвольной непрерывной функции посредством многочленов данной степени n_0 принципиально равнозначна наилучшему приближению многочлена любой более высокой степени $n > n_0$ при помощи многочленов степени n_0. Это изучение было начато Чебышевым и его учениками Золотаревым и братьями Марковыми и продолжено позднейшими последователями Чебышева, особенно в нашей стране.

Я не буду останавливаться здесь на теоремах этой новой алгебры, ставших уже классическими, и на ряде предельно точных неравенств [подобных (4)], получаемых из них.

---

[1] Заметим, что при том же условии (5) погрешность усеченного ряда Тэйлора

$$\left| f(x) - \sum_{0}^{n} \frac{f^{(k)}(c)}{k!}(x-c)^k \right| \leqslant \frac{M}{(n+1)!}\left(\frac{b-a}{2}\right)^{n+1},$$

где $c = \frac{a+b}{2}$, т. е. в 2^n раз больше (знак равенства здесь также достигается при $f^{(n+1)}(x) = M$).

Напомню лишь о некоторых основных обратимых взаимоотношениях между порядком убывания $E_n(f(x);\ a,\ b)$ при $n \to \infty$ и соответствующими свойствами функции. Согласно теореме Вейерштрасса, условие $\lim E_n(f(x);\ a,\ b) = 0$ при $n \to \infty$ может быть принято в качестве (конструктивного) определения непрерывности функции $f(x)$ в промежутке $(a,\ b)$. Условие, что $E_{n_0}(f(x);\ a,\ b) > 0$ и $E_n(f(x);\ a,\ b) = 0$ при $n > n_0$ означает, что $f(x)$ есть многочлен степени $(n_0 + 1)$.

Одним из основных фактов общей теории непрерывных функций вещественной переменной является то, что аналитические функции могут быть определены как класс функций, для которых $E_n(f(x);\ a,\ b)$ убывает в геометрической прогрессии при бесконечном возрастании n (т. е. $\overline{\lim} \sqrt[n]{E_n(f(x);\ a,\ b)} = \rho < 1$); при этом знаменатель ρ этой прогрессии тем ближе к единице, чем меньше размеры того из софокусных эллипсов, имеющих фокусами концы отрезка $(a,\ b)$, внутри которого функция $f(x)$ регулярна, в частности, $\lim \sqrt[n]{E_n(f(x);\ a,\ b)} = 0$ только в том случае, когда $f(x)$ регулярна во всей комплексной плоскости.

Таким образом, если функция $f(x)$ *неаналитическая* на $(a,\ b)$, т. е. по крайней мере в одной точке этого отрезка она не разлагается в степенной ряд, то

$$\overline{\lim} \sqrt[n]{E_n(f(x);\ a,\ b)} = 1$$

(хотя $E_n(f(x);\ a,\ b)$ стремится к нулю при $n \to \infty$ для всякой непрерывной функции). При этом закон убывания $E_n(f(x);\ a,\ b)$ весьма тесно связан с модулем непрерывности и другими дифференциальными свойствами рассматриваемой функции, а в случае дифференцируемости функции $f(x)$ наилучшее приближение тем быстрее убывает с возрастанием n, чем больше число конечных последовательных производных, которыми обладает функция. Однако, как было замечено выше, если функция $f(x)$ *неаналитическая*, то даже при существовании производных всех порядков наилучшие приближения $E_n(f(x);\ a,\ b)$ убывают к нулю медленнее, чем члены любой убывающей геометрической прогрессии. Благодаря тесной связи $E_n(f(x);\ a,\ b)$ с различными дифференциальными особенностями функции $f(x)$ удалось обнаружить некоторые глубокие зависимости между различными структурными свойствами неаналитических непрерывных функций, которые ранее оставались неизвестными. С другой стороны, возможно установить

без помощи комплексной переменной специфическое фундаментальное свойство аналитического продолжения аналитических функций, являющееся математическим выражением детерминистического принципа естествознания, как единственное их продолжение или экстраполяция за пределы отрезка (a, b), при котором сохраняется закон убывания наилучшего приближения в геометрической прогрессии. Все это, как и многие другие факты, свидетельствует о теоретической и практической целесообразности единой классификации всех непрерывных функций $f(x)$ на базе закона убывания их наилучшего приближения $E_n(f(x); a, b)$ к нулю с возрастанием n.

Точные значения всей последовательности монотонно убывающих чисел $E_n(f(x); a, b)$ обычно неизвестны и не представляют практического интереса, однако принципиальное значение имеет следующий вопрос. Пусть дана произвольная бесконечная последовательность монотонно убывающих к нулю с возрастанием n чисел $\alpha_0 \geqslant \alpha_1 \geqslant \ldots \geqslant \alpha_n \geqslant \ldots$ Существует ли функция $f(x)$ на отрезке (a, b), для которой $E_n(f(x); a, b) = \alpha_n$ при всех n?

Ответ на этот вопрос утвердителен. Но построение всех функций, имеющих данные $E_n(f(x); a, b)$, представляет большие трудности. Для случая, когда $f(x)$ многочлен (n_0+1)-й степени, т. е. $\alpha_n = 0$ при $n > n_0$, можно указать общий прием эффективного решения этой задачи. Оказывается, что, не считая произвольного постоянного слагаемого, число различных функций, т. е. многочленов (n_0+1)-й степени, для которых числа $\alpha_n = E_n(f(x); a, b) > 0$ лишь при $n \leqslant n_0$ равны 2^l, где $l \leqslant n_0+1$ есть число неравных между собой данных значений α_n при $0 \leqslant n \leqslant n_0$; а если, введя соответствующее соглашение, приписывать α_n определенный знак, то многочлен[1] определяется однозначно этими l значениями α_n. Метода для решения этой обратной проблемы теории наилучшего приближения, когда все данные числа $\alpha_n \neq 0$, пока не найдено. Метод (неопубликованный), которым я располагаю, пригоден лишь при условии, что числа α_n убывают достаточно быстро, а именно $\frac{\alpha_{n+1}}{\alpha_n} < \frac{1}{5}$. Таким образом, можно, например, утверждать, что

[1] А именно тот, который имеет погрешность $f(x) - P_n(x)$, когда она в последний раз (при возрастании x) достигает на отрезке (a, b) своего максимального значения $E_n(f(x); a, b)$.

$\frac{1}{d-x}$ является единственной (с точностью до произвольного постоянного слагаемого) абсолютно монотонной[1] на отрезке $(-1, +1)$ функцией $f(x)$, для которой при всех[2] $n \geqslant 0$

$$\alpha_n = E_n(f(x); -1, +1) = \frac{1}{(d^2-1)(d+\sqrt{d^2-1})^n} \qquad \left(d > \frac{13}{5}\right).$$

§ 6. За меру погрешности часто бывает полезно принять среднюю квадратичную погрешность

$$\int_a^b [f(x) - P_n(x)]^2 d\sigma(x) \qquad (\sigma(b) - \sigma(a) = 1),$$

где $\sigma(x)$ — некоторая данная монотонная функция, рассмотрение которой связано с известным всем практикам способом наименьших квадратов. В таком случае теория разложения функций по соответствующим ортогональным многочленам дает сравнительно простой алгорифм для вычисления многочленов, обращающих в минимум $I_n(f(x); \sigma(x))$ — указанную среднюю квадратичную погрешность. Ввиду этого весьма ценны некоторые простые неравенства,[3] связывающие $I_n(f(x); \sigma(x))$ с соответствующими наилучшими приближениями $E_n(f(x); a, b)$.

Отмечу здесь почти очевидный факт, что упомянутая выше обратная задача наилучшего приближения при этом новом мероопределении легко разрешается в самом общем случае. Какова бы ни была данная последовательность монотонно убывающих к нулю чисел $\beta_0 \geqslant \beta_1 \geqslant \ldots \ldots \geqslant \beta_n \geqslant \ldots$, все функции $f(x)$, для которых $I_n(f(x); x) = \beta_n$ (полагая, например, $\sigma(x) = x$, $a=0$, $b=1$), определяются с точностью до постоянного слагаемого равенствами

---

[1] Функция $f(x)$ называется абсолютно монотонной на отрезке, если все ее последовательные производные неотрицательны на этом отрезке.

[2] Ограничение $d > \frac{13}{5}$ (вместо $d > 1$), связанное с несовершенством метода доказательства, без сомнения, излишне.

[3] Нетрудно заметить, что $I_n(f(x)) \leqslant E^2_n(f(x))$ на любом отрезке (a, b) и при всяком $\sigma(x)$; поэтому если $f(x)$ непрерывна, то $I_n(f(x); \sigma(x)) \to 0$ при $n \to \infty$, но обратное утверждение неверно: свойство $I_n(f(x); \sigma(x)) \to 0$ равнозначно утверждению, что $\int_a^b f^2(x)\, d\sigma(x)$ имеет смысл.

$$A_n = \int_0^1 f(x)\, R_n(x)\, dx = \pm \sqrt{\beta_{n-1} - \beta_n},$$

где $R_n(x)$ — соответствующим образом нормированный многочлен Лежандра степени n на отрезке $0,1$, а A_n — коэффициенты формального разложения $f(x)$ по этим многочленам. Таким образом, все эти функции отличаются друг от друга лишь знаками коэффициентов своих разложений. В случае, когда все получаемые разложения сходятся благодаря достаточно быстрому убыванию β_n абсолютно и равномерно на $0,1$, все получаемые функции непрерывны; в противном случае среди последних могут быть и разрывные и бесконечно возрастающие функции. Вообще дифференциальная природа функций значительно менее точно характеризуется величинами $I_n(f(x); x)$, чем наилучшими приближениями $E_n(f(x); 0,1)$, что и понятно, так как среднее приближение $I_n(f(x); x)$ затушевывает локальные особенности функции.

§ 7. Вместо многочленов $P_n(x)$ для приближения любой непрерывной функции на заданном отрезке можно пользоваться и другими системами простых функций, в частности полиномами $P_n(\varphi(x))$, где $\varphi(x)$ — какая-нибудь определенная, непрерывно возрастающая на данном отрезке функция.

Полагая для определенности $a = -1$, $b = 1$, особо важное значение имеет замена $x = \cos\theta$ ($\cos\theta$ возрастает от -1 до $+1$ при изменении θ от $-\pi$ до 0). В самом деле, любую четную периодическую непрерывную функцию $F(\theta)$ с периодом 2π можно представить в виде $F(\theta) = f(\cos\theta)$, и если $P_n(x) = \sum\limits_0^n a_k x^k$ есть многочлен степени n, наименее уклоняющийся от $f(x)$ при $-1 \leqslant x \leqslant 1$, то $P_n(\cos\theta) = \sum\limits_0^n a_k \cos^k\theta = \sum\limits_0^n A_k \cos k\theta$ есть тригонометрическая сумма порядка n, наименее уклоняющаяся от $F(\theta)$. Таким образом, если обозначить через $E_n^*(F(\theta))$ наилучшее приближение периодической функции $F(\theta)$ посредством тригонометрической суммы порядка n, то $E_n^*(F(\theta)) = E_n^*(f(\cos\theta)) = E_n(f(x); -1,1)$. Отсюда видно, что проблема приближения любых непрерывных функций многочленами по существу эквивалентна

проблеме приближения периодических функций тригонометрическими суммами, и теоремами, аналогичными ранее упомянутым, регулируется тесная связь между дифференциальной природой периодической функции $F(\theta)$ и законом убывания ее наилучшего тригонометрического приближения $E_n^*(F(\theta))$.

Любопытно, что этот совершенно очевидный теперь факт был замечен только в начале нашего столетия, и, например, еще в 1905 г. один из крупнейших основателей современной теории функций — Борель отмечал в одной из своих известных монографий, что, в то время как тригонометрический ряд Фурье $f(\theta) = \sum A_k \cos k\theta + B_k \sin k\theta$ для бесконечно дифференцируемой периодической функции бесконечно дифференцируем, разложения по многочленам, обладающего тем же свойством, еще не известно; а между тем из указанной выше замены переменной видно, что требуемым свойством обладает разложение по многочленам Чебышева $\cos k \arccos x$.

Вообще многочисленные результаты глубоких исследований прошлого столетия о сходимости тригонометрических рядов, выросших на почве бесконечных формальных процессов интегрирования линейных уравнений математической физики, естественно влились теперь в общую конструктивную теорию функций, а теоремы новой экстремальной алгебры Чебышева существенно содействовали дальнейшим успехам теории тригонометрических рядов.

§ 8. Если первоначальная концепция аналитической функции неразрывно связана с комплексной переменной, то тригонометрические ряды, несмотря на их способность представлять любую функцию в промежутке не более периода, всегда выражают некоторую периодическую функцию на *всей* вещественной оси.

До недавнего времени периодические функции и функции, стремящиеся к определенным конечным пределам [1] при $x = \pm\infty$, были един-

[1] Если $f(\pm\infty) = 0$, то $f(x)$ может быть, например, равномерно приближена рациональными дробями $\sum_1^n \frac{A_k + B_k x}{(1+x^2)^k}$; если же $f(-\infty) = A$, $f(+\infty) = B$, то достаточно заменить $f(x)$ через $f(x) - \frac{A + Be^x}{1+e^x}$, чтобы свести этот случай к предыдущему.

ственными равномерно непрерывными на всей оси функциями, представление которых равномерно сходящимися рядами при всех x $(-\infty < x < \infty)$ было известно. Между тем, решение этой проблемы имеет существенное значение для анализа, в частности для вещественной теории обыкновенных дифференциальных уравнений.

Пусть задана, например, система дифференциальных уравнений

$$\frac{dx_i}{dt} = P_i(x_1, \ldots, x_n) \qquad (i=1, 2, \ldots, n), \tag{6}$$

где P_i — при всех конечных значениях $x_1, x_2, \ldots, x_n$ непрерывные функции координат $x_1, x_2, \ldots, x_n$ точки, движущейся в n-мерном пространстве. Рассматривая $x_k(t)$ как функции времени t, важно исследовать всевозможные траектории точки (устойчивые в смысле Пуассона), на которых все координаты $x_k(t)$ *ограничены* при любом t $(-\infty < t < \infty)$. Из уравнения (6) видно, что все $x_k(t)$, соответствующие каждой данной траектории такого рода, удовлетворяют некоторому, не зависящему от t условию Липшица, т. е. *равномерно* непрерывны [1].

В некоторых интересных случаях эти функции оказываются периодическими или почти-периодическими, т. е. могут быть равномерно приближены при всех t $(-\infty < t < \infty)$ суммами конечного числа слагаемых вида

$$F_p(t) = \sum_0^p A_k \cos \lambda_k (t - t_k), \tag{7}$$

где A_k, t_k, λ_k $(0 \leqslant \lambda_k < p)$ — любые вещественные числа, при p достаточно большом. Теория почти-периодических функций, представляющих существенное обобщение периодических функций, которые соответствуют частному случаю, когда все числа λ_k соизмеримы, — одно из крупнейших достижений анализа последнего времени. Но почти-периодические функции, конструктивно определяемые (подобно непрерывным перио-

[1] Координаты x_k, рассматриваемые как функции длины дуги $s = \int_0^t \sqrt{\Sigma\ dx_i^2}$, отсчитываемой на любой траектории от произвольно взятой на ней обыкновенной точки (где $\Sigma P_i^2 > 0$), являются равномерно непрерывными (с модулем непрерывности $\omega(h) = h$) при $-\infty < s < \infty$ даже и на тех кривых, которые удаляются в бесконечность).

9*

дическим функциям) условием, что их наилучшее приближение тригонометрическими суммами (7) стремится к нулю при $p \to \infty$, далеко не исчерпывают всей совокупности равномерно непрерывных на всей вещественной оси функций. Для этой цели необходимо использование более обширного класса так называемых целых функций *конечной степени*, включающего суммы (7) как частный случай. Мы называем целую функцию

$$f(x) = \sum_{0}^{\infty} a_k \frac{x_k}{k!}$$

функцией конечной степени p, если[1] $\overline{\lim} \sqrt[k]{|a_k|} = p$. Легко проверить, что суммы (7) действительно удовлетворяют этому условию.

В частности, $\cos(x - a)$ есть целая функция первой степени, характеризуемая замечательным экстремальным свойством:

Если целая функция $f(x)$ степени p удовлетворяет при всех x $(-\infty < x < \infty)$ неравенству $|f(x)| \leqslant 1$, то

$$|f^{(k)}(x)| \leqslant p^k \qquad (-\infty < x < \infty) \tag{8}$$

для всех целых k; при этом если знак равенства в (8) имеет место для какого-нибудь $k > 0$, то он сохраняется для всех k, и единственными функциями, для которых равенство фактически может осуществляться, являются элементарные функции $\cos p(x - a)$.

Благодаря этой теореме, класс (B_p) ограниченных (при всех x) функций степени p может быть охарактеризован (без предварительного введения понятия целой функции) следующим экстремальным свойством: если $f(x) \in (B_p)$ и $|f(x)| \leqslant M$, то $f'(x) \in (B_p)$ и $|f'(x)| \leqslant Mp$ при всех x.

Из этой же теоремы выводится, что если $A_p(f(x))$ означает наилучшее приближение функции $f(x)$, ограниченной на всей оси при помощи целых функций степени p, то равенство $A_p(f(x)) \to 0$ при $p \to \infty$ является необходимым и достаточным признаком равномерной непрерывности функции $f(x)$ на всей оси. Таким образом, в частности, все устойчивые в смысле Пуассона решения уравнений (6) обладают свойством, что $A_p(f(x)) \to 0$ при $p \to \infty$.

[1] Если $f_1(x)$ есть целая функция первой степени, то $f_1(px)$ есть при любом $p > 0$ целая функция степени p и $[f_1 x]^n$ есть функция степени n при всяком целом n.

Можно показать, что если функция $f(x)$ периодическая, то $A_p(f(x)) = E_p^*(f(x))$. Кроме того, закон убывания $A_p(f(x))$ в общем случае любых равномерно непрерывных функций так же точно характеризует дифференциальные свойства функции на всей оси, как и в случае периодичности, т. е. взаимная связь здесь та же, что и между соответствующими дифференциальными свойствами функции $f(x)$, непрерывной на любом данном отрезке (a, b), и ее наилучшим приближением $E_n(f(x);\ a, b)$ при помощи многочленов степени n.

Теория приближения функций посредством целых функций конечной степени не только служит естественным обобщением приближения периодических и почти-периодических функций посредством тригонометрических сумм (7), но является также необходимым продолжением теории наилучшего приближения функций многочленами, как это видно из следующих предельных неравенств:[1]

$$A_{p_1}(f(x)) \leqslant \lim_{n=\infty} E_n(f(x);\ -\lambda,\ \lambda) \leqslant A_p(f(x)),$$

каковы бы ни были $p_1 < p_0 = \lim \frac{n}{\lambda} < p$. Таким образом, если p_0 есть точка непрерывности функции $A_p(f(x))$, монотонно убывающей при возрастании p, то

$$\lim_{n=\infty} E_n\left(f(x);\ -\frac{n}{p_0},\quad \frac{n}{p_0}\right) = A_{p_0}(f(x)).$$

Заканчивая эту статью, которая имела целью показать, что экстремальные неравенства и функции или классы функций, характеризуемые простыми экстремальными свойствами, служат путеводной нитью, направляющей плодотворное развитие анализа, я считаю не лишним отметить, что для иллюстрации этого положения можно было бы также воспользоваться многими другими не менее яркими примерами. В создании и развитии теорий, о которых здесь была речь, так же как и в других важных областях математики, значительное участие принимали советские математики.

---

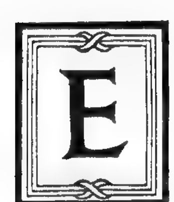
[1] Условие ограниченности $f(x)$ может быть заменено значительно более общими предположениями.

Член-корреспондент АН СССР
П. С. Александров

ТЕОРЕМЫ ДВОЙСТВЕННОСТИ В КОМБИНАТОРНОЙ ТОПОЛОГИИ

§ 1

ТЕОРЕМА ЖОРДАНА



сли S окружность, лежащая в плоскости R, то всякий непрерывный путь, соединяющий какую-либо точку, лежащую вне окружности S, с какою-либо точкой, лежащей внутри окружности S, непременно пересекается с окружностью S. Другими словами, окружность S *разбивает плоскость на две области*: внешнюю (к этой окружности) и внутреннюю. Тем же свойством обладает и всякая замкнутая ломаная линия без кратных точек. Свойство окружности разбивать плоскость на две области непосредственно вытекает из определения окружности и простейших теорем о связных множествах. То же свойство в применении к замкнутым ломаным линиям требует для своего доказательства некоторых теорем элементарной геометрии. Легко доказать при помощи простых предложений дифференциальной геометрии, что и всякая замкнутая линия на плоскости, имеющая касательную в каждой точке какой-либо (хотя бы и очень малой) своей дуги, также разбивает плоскость на две области, внешнюю и внутреннюю.

Но как быть с замкнутыми кривыми, не имеющими касательной ни в какой своей точке?

Можно отнестись к ним по-разному. Можно, во-первых, подобно Эрмиту, объявить эти кривые (вместе с лежащими в их основе непрерывными функциями, ни в одной точке не имеющими производной) «плачевной язвой» на теле математики и более ими не интересоваться; но можно также, определив сначала как следует понятие простой (т. е. не имеющей кратных точек) замкнутой непрерывной кривой[1], попытаться доказать, что всякая такая кривая разбивает плоскость. Этот второй путь был избран другим знаменитым французским математиком — Жорданом, впервые доказавшим (в своем курсе анализа) только-что сформулированное предложение, известное теперь под именем теоремы Жордана и ставшее одним из весьма знаменитых предложений современной математики. Почему же эта теорема, очевидная в элементарных случаях и казавшаяся некоторым математикам конца прошлого века излишней в том единственном случае, когда ее доказательство представляет действительные трудности, выдержала полувековое испытание и ныне всеми признается замечательной и интересной теоремой?

Во-первых, потому, что без нее невозможно обойтись в целом ряде новейших исследований в такой, например, «классической» области математики, как теория функций комплексного переменного; во-вторых, потому, что доказательство теоремы Жордана любым из многочисленных предложенных для этой цели методов (они были предложены Вебленом, Александером, Брауэром, Лебегом, де-ла-Валле Пуссеном, Эрхардом Шмидтом, Антуаном, мною и многими другими) представляет большие трудности, лежащие в существе дела и преодолеваемые лишь той или иной сильной математической конструкцией; в-третьих, потому, что в топологических вопросах (а к этому кругу вопросов и относится теорема Жордана) вещи, на первый взгляд кажущиеся очевидными, часто оказываются неверными и поэтому всегда требуют тщательного исследования. В случае теоремы Жордана такое тщательное исследование геометрической обстановки подтвердило первоначальное чувство очевидности. В других случаях результат получался противоположный: наша непосредственная геометрическая интуиция оказывалась неверной (примеры этого мы скоро увидим). Не удивительно поэтому, что к доказательству или опровержению каждой правдоподобной и в то же время достаточно про-

[1] Такое определение дано в § 3 настоящей статьи.

стой геометрической гипотезы математики стали относиться с интересом, так как заранее не известно, какова будет судьба этой гипотезы. Топология, в отличие от многих других, менее молодых частей математики, имеет как-раз то преимущество, что в ней и до сих пор высказываются, доказываются и опровергаются гипотезы, формулировки которых обладают элементарной убедительностью, доступной всякому студенту-математику, но исследование которых сталкивается с трудностями, часто очень большими, а иногда при современном состоянии науки и непреодолимыми. Как бы то ни было, теорема Жордана может служить типичным примером глубокой и сильной теоремы, являющейся так называемым «чистым обобщением», т. е. весь интерес которой зависит именно от того, что она утверждается в применении к достаточно широкому классу объектов (именно, к классу кривых, соответствующих классу непрерывных функций, нигде не дифференцируемых).

§ 2

ПРОСТЕЙШИЕ ТОПОЛОГИЧЕСКИЕ ПОНЯТИЯ ЛЕММА ШПЕРНЕРА

Предположим, что A и B два каких-нибудь множества и что каждому элементу x множества A поставлен в соответствие определенный элемент $y = f(x)$ множества B. Тогда говорят, что дано отображение f множества A во множество B, или что дана функция $y = f(x)$, определенная на множестве A и имеющая своими значениями элементы множества B (если и A и B являются множествами вещественных или комплексных чисел, то приходим к обычному определению функции в общем смысле этого слова от вещественного или комплексного переменного, с вещественными или комплексными значениями). Это общее определение отображения или функции может приводить к интересным вопросам даже в том, казалось бы, совершенно тривиальном случае, когда и A и B состоят лишь из конечного числа элементов. Рассмотрим, например, треугольник $e_0\ e_1\ e_2$ и его подразделение K на более мелкие треугольники. Через A мы обозначим множество всех вершин, возникших при подразделении K треугольника $e_0\ e_1\ e_2$ (включая, конечно, и сами вер-

шины e_0, e_1, e_2 первоначального треугольника), а через B обозначим множество, состоящее из трех вершин e_0, e_1, e_2. Предположим теперь, что каждому элементу x множества A (т. е. каждой вершине подразделения K) поставлен в соответствие некоторый элемент $f(x)$ множества B, т. е одна из вершин e_0, e_1, e_2, так что соблюдаются два следующих условия:

1°. Каждой вершине e_0, e_1, e_2 (рассматриваемой как один из элементов множества A) ставится в соответствие сама эта вершина (являющаяся ведь элементом и множества B):

$$f(e_0)=e_0, \qquad f(e_1)=e_1, \qquad f(e_2)=e_2.$$

2°. Если данная вершина x, принадлежащая множеству A (т. е. являющаяся одной из вершин подразделения K), лежит на какой-нибудь стороне нашего треугольника, например на стороне $e_0\, e_1$, то поставленная ей в соответствие вершина $f(x)$ должна быть одним из концов той стороны треугольника $e_0\, e_1\, e_2$, на которой лежит вершина x (т. е. в нашем случае, когда x лежит на $e_0\, e_1$, вершина $f(x)$ есть либо e_0 либо e_1). Больше никаких ограничений на отображение f не налагается. Оказывается, что при этих условиях *в подразделении K непременно найдется хотя бы один треугольник T, трем вершинам которого поставлены в соответствие все три вершины e_0, e_1, e_2 основного треугольника* (т. е. никаким двум вершинам треугольника T не поставлена в соответствие одна и та же вершина треугольника $e_0\, e_1\, e_2$).

Это утверждение, которое можно объяснить любому интересующемуся математикой школьнику 8—9 класса, допускает немедленное обобщение на случай произвольного тетраэдра (треугольной пирамиды), а также на случай тетраэдров четырех и более измерений (т. е. на случай так называемых *n-мерных симплексов*, являющихся простейшими многомерными выпуклыми многогранниками) и известно под названием теоремы или *леммы Шпернера*, по имени математика, доказавшего его лет двадцать тому назад. Само доказательство — совершенно элементарно и по существу также доступно школьнику 8—9 класса; тем не менее я не думаю, чтобы оно удалось даже всем тем из моих читателей, которые захотели бы это доказательство провести. Насколько глубок геометрический факт, выражаемый леммой Шпернера, мы увидим на протяжении этого же параграфа, но предварительно продолжим изложение простейших понятий, связанных с понятием отображения или функции.

Пусть множества A и B являются точечными множествами, т. е. все их элементы являются точками, скажем, m-мерного эвклидова пространства (или, если угодно, бесконечномерного гильбертова пространства, хорошо известного физикам-теоретикам). Тогда можно говорить о расстоянии между любыми двумя точками каждого из множеств A и B. Отображение f множества A во множество B называется *непрерывным в данной точке* x множества A если, каково бы ни было положительное число ε, можно подобрать такое положительное число δ, что для каждой точки x множества A, расстояние которой от точки x_0 меньше, чем δ, расстояние точки $f(x)$ от точки $f(x_0)$ оказывается меньше ε. Если отображение f непрерывно во всякой точке множества A, то оно называется непрерывным отображением множества A во множество B. В частном случае, когда A и B суть множества действительных или комплексных чисел, мы получаем известное из элементов анализа определение непрерывной функции. Заметим кстати: если $y=f(x)$ есть определенная, скажем, на отрезке $0 \leqslant x \leqslant 1$ непрерывная функция, то график этой функции есть множество Φ, состоящее из всех точек плоскости, имеющих вид $x, f(x))$, где $0 \leqslant x \leqslant 1$. Ставя в соответствие каждой точке x отрезка $0 \leqslant x \leqslant 1$ соответствующую точку $(x, f(x))$ множества Φ, получим отображение φ отрезка $0 \leqslant x \leqslant 1$ на множество Φ, и отображение это будет очевидно непрерывным. Но отображение φ будет обладать еще и свойством *взаимной однозначности*: это значит, что каждая точка $(x, f(x))$ множества Φ является образом одной и только одной точки отрезка $[0;1]$, именно той точки x, которая является абсциссой данной точки $(x, f(x))$ множества Φ. Если дано взаимно однозначное отображение φ какого-нибудь множества A на какое-нибудь множество B, то, ставя в соответствие каждой точке y множества B ту единственную точку x множества A, для которой $\varphi(x)=y$, получим отображение множества B на множество A, *обратное* отображению φ и обозначаемое обычно через φ^{-1}.

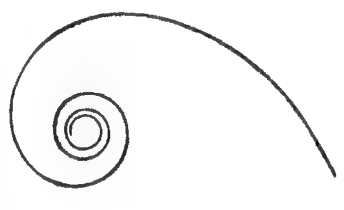
Пусть φ какое-либо взаимно однозначное отображение данного точечного множества A на точечное множество B. *Если не только отображение* φ, *но и обратное ему отображение* φ^{-1} *непрерывны, то отображение* φ *называется топологическим отображением* (*или гомеоморфизмом* [1]); в этом

[1] Отображение, обратное к взаимно однозначному и непрерывному отображению φ, может не быть непрерывным: достаточно взять за A полуинтервал $0 \leqslant x < 2\pi$ числовой

случае отображение φ^{-1}, конечно, тоже будет топологическим. Упомянутое нами отображение φ отрезка $0 \leqslant x \leqslant 1$ на множество Φ (являющееся графиком какой-либо непрерывной на $[0; 1]$ функции $f(x)$) может служить примером топологического отображения.

Два точечных множества называются *гомеоморфными* или топологически эквивалентными между собою, если одно из них можно топологически отобразить на другое.

Мы только что видели, что график любой функции, непрерывной на замкнутом отрезке, гомеоморфен этому отрезку. Множества, гомеоморфные замкнутому отрезку, называются *простыми дугами*. Таким образом, график всякой функции $y=f(x)$, непрерывной на каком-либо замкнутом отрезке, является простой дугой. Однако не всякая простая дуга, лежащая на плоскости, может быть представлена как график непрерывной функции в декартовых координатах. Например, как бы мы ни выбирали на плоскости оси прямоугольных (или даже аффинных) координат, мы никогда не сможем представить в виде графика функции $y=f(x)$ кривую, даваемую в полярных координатах уравнением



Фиг. 1

$$r = \frac{1}{\varphi}$$

(«гиперболическая спираль»), пополненную полюсом $r=0$ (фиг. 1), хотя эта кривая и является простой дугой.

Некоторое свойство точечного множества называется *топологическим* или топологически-инвариантным, если, принадлежа данному множеству,

---

прямой, а за B окружность $x^2 + y^2 = 1$ и определить φ как отображение, ставящее в соответствие точке x, $0 \leqslant x < 2\pi$, точку окружности B, имеющую полярный угол x (отображение заключается таким образом в «навертывании» полуинтервала $0 \leqslant x < 2\pi$ на окружность B); отображение φ взаимно однозначно (так как точка 2π во множество A не входит); если $b_0 = (1; 0)$ есть точка окружности, являющаяся началом отсчета углов (т. е. точка с полярным углом 0), то в b_0 отображение φ^{-1} не будет непрерывным (так как образ точки b_0 при отображении φ^{-1} есть точка 0 отрезка $0 \leqslant x < 2\pi$ и в любой близости точки b_0 имеются точки у нашей окружности, для которых $f^{-1}(y)$ сколь угодно мало отличается от 2π).

это свойство принадлежит и всем множествам, гомеоморфным данному.[1] Так, все простые дуги имеют одни и те же топологические свойства; с точки зрения топологии совершенно несущественно, является ли данная простая дуга прямолинейным отрезком, или дугой окружности, или гиперболической спиралью, изображенной на фиг. 1, или кривой, ни в какой своей точке не имеющей касательной.

Если нам дано какое-нибудь точечное множество, то относительно некоторых его свойств сразу можно сказать, что они являются топологически-инвариантными. К числу таких свойств принадлежит, например, свойство *компактности*, заключающееся в том, что из каждой бесконечной последовательности точек данного множества A можно выделить подпоследовательность, сходящуюся к какой-нибудь точке этого же множества A; *топологическая инвариантность* этого свойства непосредственно вытекает из того, что при непрерывном отображении f множества A на множество B всякая последовательность точек A, сходящаяся к некоторой точке a множества A, переходит в последовательность точек B, сходящуюся к точке $b = f(a)$ множества B.

С другой стороны, существуют свойства точечных множеств, топологическая инвариантность которых требует более или менее сложных доказательств. Чтобы получить пример такого свойства, вспомним, что множество A (лежащее в эвклидовом пространстве любого числа измерений или в гильбертовом пространстве) называется замкнутым (в том пространстве, точечным множеством которого множество A является), если оно содержит все свои предельные точки (точка x_0, принадлежащая

[1] При этом, когда мы говорим «все множества, гомеоморфные данному», то разумеем все точечные множества (т. е. все множества, лежащие в каком-либо эвклидовом или в гильбертовом пространстве). Иногда говорят о топологической инвариантности какого-либо свойства данного множества в *пределах некоторого, заранее очерченного класса множеств*: например, одним из глубоких топологических фактов является следующая, доказанная Брауэром теорема, известная под названием теоремы об инвариантности открытых множеств. Если множество A есть открытое (т. е. дополнительное к замкнутому) множество эвклидова m-мерного пространства R^m и *если лежащее в* R^m *множество* B гомеоморфно множеству A, то B также открыто в R^m. Другими словами, свойство множества, *лежащего в m-мерном эвклидовом пространстве*, быть открытым множеством пространства R^m есть топологически инвариантное свойство (*в пределах класса множеств, лежащих в m-мерном эвклидовом пространстве*).

или нет множеству A, называется предельной точкой этого множества, если существует последовательность отличных от точки x_0 точек множества A, сходящаяся к точке x_0). Множество всех точек какой-либо прямой замкнуто (на плоскости) или в эвклидовом пространстве и в то же время гомеоморфно открытому интервалу: для того чтобы убедиться в этом, достаточно заметить, что функция $y = \operatorname{tg} x$, рассматриваемая на интервале $-\frac{\pi}{2} < x < \frac{\pi}{2}$, дает нам топологическое отображение этого интервала на всю прямую $-\infty < y < +\infty$. Так как интервал не является замкнутым множеством, то мы видим, что свойство точечного множества A быть замкнутым (в содержащем его эвклидовом пространстве) не есть топологическое свойство этого множества. Тот же пример прямой и гомеоморфного ей интервала учит, что неограниченное множество в эвклидовом (или гильбертовом) пространстве может быть гомеоморфно ограниченному. Таким образом, ни замкнутость, ни ограниченность множества не являются топологическими свойствами этого множества. Между тем можно легко доказать, что множество, лежащее в эвклидовом пространстве, тогда и только тогда является компактным в себе, когда оно одновременно и замкнуто и ограниченно. Таким образом, совокупность двух свойств — замкнутости и ограниченности для множеств, лежащих в эвклидовом пространстве, оказывается топологически инвариантным (а именно, эквивалентным свойству компактности в себе), хотя ни одно из этих свойств, взятое отдельно, топологически инвариантным не является. Точечные множества, компактные в себе, играют в современной топологии особенно важную роль; они для краткости называются *компактами*.

То рассуждение, с помощью которого мы убедились в топологической инвариантности понятия компактности, позволяет доказать и более сильное утверждение, а именно, что *всякое точечное множество, являющееся образом некоторого компакта при некотором непрерывном отображении, само является компактом*. Другими словами, свойство компактности сохраняется не только при топологических, но даже при всех непрерывных отображениях. Кроме того, легко доказать, что всякое взаимно однозначное и непрерывное (в одну сторону) отображение компакта есть топологическое отображение (т. е. обратное отображение для такого отображения будет также непрерывным).

§ 3

О РАЗБИЕНИИ ЭВКЛИДОВА ПРОСТРАНСТВА ЛЕЖАЩИМИ В НЕМ ЗАМКНУТЫМИ МНОЖЕСТВАМИ, О МАЛЫХ СДВИГАХ ЭТИХ МНОЖЕСТВ И О ПОНЯТИИ РАЗМЕРНОСТИ

После того как мы ввели простейшие топологические понятия, мы можем вернуться к тем двум теоремам, о которых была речь в § 1 и в начале § 2: к теореме Жордана и к так называемой лемме Шпернера.

Прежде всего мы можем теперь определить наиболее общую замкнутую непрерывную кривую *как точечное множество, гомеоморфное окружности,* после чего теорема Жордана получает такую формулировку:

Теорема Жордана. Всякая замкнутая непрерывная кривая, лежащая на плоскости, разбивает плоскость на две области и является их совместной границей. Другими словами, свойство окружности разбивать плоскость на две области (и быть общей границей этих двух областей) есть, в пределах класса всех плоских множеств, топологически инвариантное свойство.

Как уже упоминалось, теорема Жордана и сейчас, несмотря на весь прогресс топологии за последние десятилетия, остается довольно трудной теоремой: я не знаю никакого доказательства этой теоремы, полное изложение которого (рассчитанное на лиц, владеющих математикой в размере обычного университетского курса) умещалось бы менее чем в две двухчасовые лекции. Попытаемся понять, в чем здесь трудность. Если C — простая замкнутая линия, то существует топологическое отображение f окружности S^1 на кривую C. Вообразим на минуту, что топологическое отображение f окружности S^1 на кривую C можно распространить на всю плоскость R^2, т. е. построить такое топологическое отображение F плоскости R^2 на себя, которое в точках окружности S^1 совпадает с отображением f. Очевидно, что при отображении F две области, на которые окружность S^1 разбивает плоскость, перейдут в две области, на которые плоскость разобьется кривой C, и C будет совместной границей этих двух областей. Теорема Жордана будет доказана. *Однако a priori не ясно, что топологическое отображение f окружности S^1 на*

данную замкнутую линию можно распространить на всю плоскость, и доказать это предложение представляло очень большие трудности.

Поэтому при доказательстве теоремы Жордана стали искать другие пути, и они были найдены, разными авторами — разные. Наконец, только около 1920 г. французский математик Антуан доказал, что всякое топологическое отображение окружности (а также всякое топологическое отображение прямолинейного отрезка) на какую-либо плоскую кривую всегда можно распространить на всю плоскость. Этот факт сам по себе очень интересен и, конечно, вполне оправдает трудности его доказательства, тем более что топологическое отображение окружности (и даже, как показал тот же Антуан, отрезка) на какую-либо кривую, лежащую в трехмерном пространстве, вообще говоря, нельзя распространить на всё трехмерное пространство.

Таким образом, доказанная Антуаном теорема о распространении есть по своему существу теорема топологии плоскости, не поддающаяся обобщению на пространство более двух измерений. Если же ограничиться самой теоремой Жордана, то — как метод ее доказательства — теорема Антуана о распространении топологических отображений окружности не заслуживает предпочтения перед другими, прямыми методами (предложенными Брауэром, Александером и др.).

Доказательство теоремы Жордана естественно выдвигает ряд вопросов:

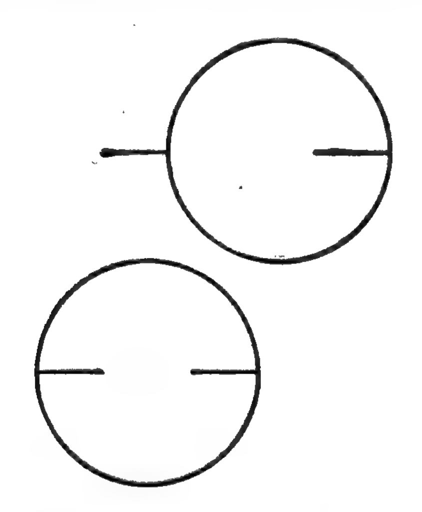
1) возникает вопрос о топологической инвариантности (в пределах класса всех плоских замкнутых множеств) свойства разбивать плоскость на заданное число областей;

2) естественно пытаться доказать так называемую n-мерную теорему Жордана, утверждающую, что всякое n-мерное замкнутое многообразие (например, всякое n-мерное сферическое многообразие, т. е. топологический образ n-мерной сферы), лежащее в $(n+1)$-мерном эвклидовом пространстве, разбивает пространство на две области и является их совместной границей;

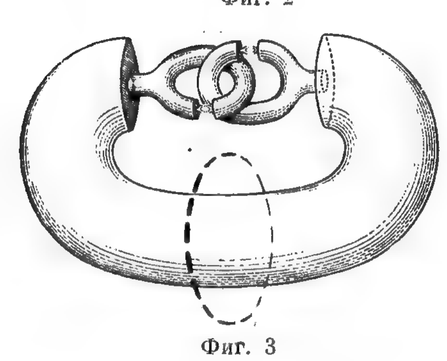
3) можно ставить вопрос о разбиении пространства *во всей общности*, т. е. пытаться доказать, что *два гомеоморфных замкнутых множества, лежащие в n-мерном эвклидовом пространстве, всегда разбивают это пространство на одно и то же число областей.*

Заметим, что ни одна из этих задач не может быть решена тем мето-

дом «распространения отображений», которым Антуан доказал обычную теорему Жордана для плоскости: две фигуры, изображенные на фиг. 2, очевидно, гомеоморфны между собою, но никакое топологическое отображение одной из них на другую не может быть распространено на всю плоскость.



Фиг. 2



Фиг. 3

Гораздо сложнее построить пример поверхности, лежащей в трехмерном пространстве и гомеоморфной сфере, которая, однако, обладает тем свойством, что никакое топологическое отображение этой поверхности на сферу не может быть распространено на все трехмерное пространство. Пример такой поверхности был впервые построен Антуаном; пример Антуана, несколько упрощенный Александером, изображен на фиг. 3.

Поверхность Антуана — Александера разбивает пространство на две области — внутреннюю (т. е. ограниченную) и внешнюю (уходящую в бесконечность). Однако *внешняя область к этой поверхности не гомеоморфна внешней области к шару* (хотя сама поверхность, как читатель помнит, гомеоморфна сфере, т. е. обыкновенной шаровой поверхности). В самом деле, во внешней области к нашей поверхности можно найти самую обыкновенную замкнутую кривую (например, окружность, изображенную на фиг. 3 пунктиром), которая никакой непрерывной деформацией (протекающей во внешней области к нашей поверхности) не может быть стянута в одну точку. Из существования во внешней области к поверхности Антуана таких замкнутых кривых, очевидно, вытекает, что

эта внешняя область не гомеоморфна внешней области к шару, откуда в свою очередь следует, что топологическое отображение шаровой поверхности на поверхность Антуана — Александера не может быть распространено на все пространство (оно не может быть даже распространено ни на какую сколь угодно малую окрестность каждой из двух поверхностей, так как не стягиваемые в точку кривые можно найти в любой близости к поверхности Антуана — Александера).

Нетрудно построить поверхность, гомеоморфную сфере и обладающую тем свойством, что в ее *внутренней* области будет бесконечно много замкнутых кривых, не стягиваемых в точку; достаточно подвергнуть пространство с лежащей в нем поверхностью Антуана — Александера преобразованию взаимными радиусами-векторами.

Первым доказательством n-мерной теоремы Жордана мы обязаны Брауэру (почему эта теорема и называется обычно теоремой Жордана — Брауэра). Теорему о топологической инвариантности числа областей, на которые плоское замкнутое множество разбивает плоскость, также впервые доказал Брауэр в 1913 г.

Общая теорема о том, что два гомеоморфных замкнутых множества разбивают n-мерное пространство на одно и то же число областей, была впервые доказана мною; почти одновременно со мною тот же результат был другим методом получен независимо друг от друга Лефшецем и Франклем. Все эти результаты относятся к 1927 г.

Метод, которым я шел в своем первом доказательстве теоремы об инвариантности числа областей, до сих пор остается самым простым (если интересоваться именно задачей о разбиении пространства, а не ее дальнейшими обобщениями); полное изложение доказательства, не предполагающее никаких предварительных знаний в области топологии, может быть сделано в пять двухчасовых лекций. Метод этот, нашедший многочисленные и разнообразные применения в различных топологических вопросах, есть метод так называемых *«малых» сдвигов* (или ε-сдвигов), переводящих замкнутые множества в полиэдры; я вкратце изложу, в чем здесь дело.

Пусть ε — положительное число и пусть Φ — компакт, лежащий, как всегда, в эвклидовом m-мерном (или в гильбертовом) пространстве R; таким образом, в случае эвклидова пространства R^m наше Φ — просто замкнутое ограниченное множество этого пространства.

Пусть дано непрерывное отображение f множества Φ на множество Φ', лежащее в том же пространстве R, что и множество Φ; если при этом расстояние каждой точки x множества Φ от ее образа $f(x)$ меньше ε, то непрерывное отображение f называется *ε-сдвигом* множества Φ (в пространстве R). Важность ε-сдвигов основана на том, что при их помощи можно, сколь угодно мало видоизменяя множество Φ (т. е. именно подвергая его ε-сдвигу при произвольно малом ε), превратить его во множество весьма простой и обозримой структуры, а именно в *полиэдр*; при этом под полиэдром мы понимаем сумму конечного числа *выпуклых многогранников* произвольного числа измерений, т. е. по существу — фигуру элементарной многомерной геометрии.

Центральное место в моем доказательстве занимает следующее предложение, представляющее и самостоятельный интерес:

Если Φ разбивает m-мерное пространство R^m на k областей (где k — конечное число), то, как бы ни было мало $\varepsilon > 0$, можно путем ε-сдвига перевести Φ в полиэдр, разбивающий пространство R^m также на k областей, причем при достаточно малом $\varepsilon > 0$ множество Φ невозможно перевести в полиэдр, разбивающий пространство меньше чем на k областей. Если Φ разбивает пространство на бесконечное число областей, то каково бы ни было натуральное число k, всякий полиэдр, в который при достаточно малом ε переходит путем ε-сдвига множество Φ, разбивает пространство на число областей, большее, чем k.

После того как доказана эта центральная теорема, доказывается без большого труда следующее предложение:

Пусть в m-мерном эвклидовом пространстве R^m даны два гомеоморфных между собою замкнутых ограниченных множества. Если при любом $\varepsilon > 0$ одно из этих множеств может быть посредством ε-сдвига переведено в полиэдр, разбивающий пространство на k областей, то то же верно и для второго множества.

Из только что сформулированных предложений следует уже без всякого труда, что два гомеоморфных замкнутых ограниченных множества разбивают m-мерное пространство на одно и то же (конечное или бесконечное) число областей. Таким образом, один из важнейших топологических инвариантов замкнутых множеств — число областей, на которые эти множества разбивают пространство, оказывается возможным свести к элементарно комбинаторным свойствам тех полиэдров, в кото-

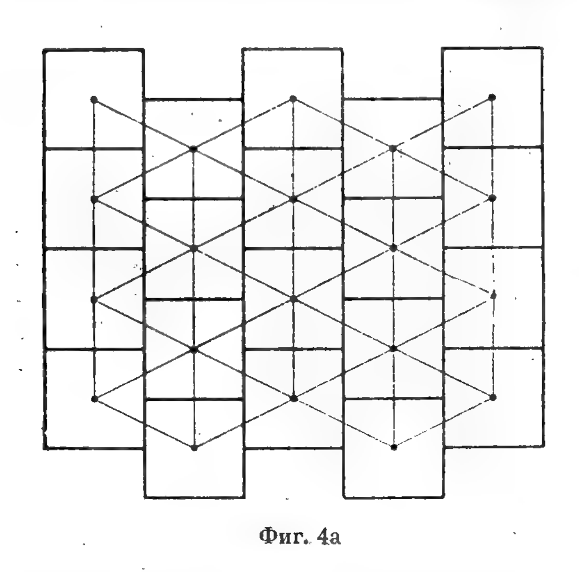
рые посредством сколь угодно малых сдвигов можно перевести данные замкнутые множества. Рассмотрим поближе эти полиэдры. Для этого нам необходимо одно построение, которое лежит в основе очень большого числа современных исследований общих вопросов топологии и которое было введено мною еще в 1925 г. Речь идет о построении так называемого *нерва* данной конечной системы множеств

$$E_1, \ldots, E_s, \tag{1}$$

лежащих в каком-нибудь пространстве R. Каждому множеству E_i нашей системы отнесем точку e_i, которую для простоты берем в том же пространстве R, в котором лежат рассматриваемые нами множества (1) (в большинстве случаев оказывается целесообразным брать точку e_i достаточно близко к соответствующему множеству E_i). Рассмотрим теперь всевозможные пары множеств E_i, E_j системы (1). Если пересечение $E_i E_j$ (т. е. множество всех точек, принадлежащих одновременно обоим множествам E_i и E_j) оказывается непустым, мы соединяем точки e_i и e_j отрезком $|e_i e_j|$; если же $E_i E_j$ пусто (т. е. если у множеств E_i и E_j нет ни одной общей точки), то мы отрезка $|e_i e_j|$ не проводим. После того как мы провели все отрезки $|e_i e_j|$, соответствующие пересекающимся парам множеств E_i, E_j, переходим к рассмотрению всех троек множеств E_i, E_j, E_k: в том и только в том случае, когда имеется хотя бы одна точка, принадлежащая всем трем множествам E_i, E_j, E_k (т. е. когда пересечение $E_i E_j E_k$ этих множеств непусто), строим треугольник с вершинами e_i, e_j, e_k (в этом случае три стороны этого треугольника уже построены, так как пересечения $E_i E_j$, $E_j E_k$, $E_i E_k$ тем более непусты). Таким же точно образом мы четверкам множеств E_i, E_j, E_k, E_l с непустым пересечением заставляем соответствовать *трехмерные симплексы* (т. е. тетраэдры) $|e_i, e_j, e_k, e_l|$ и т. д. Построение кончается на n-мерных симплексах, где $n+1$ есть *кратность* системы множеств (1), т. е. наибольшее число элементов этой системы, имеющих непустое пересечение. В результате нашего построения мы получаем конечное множество симплексов различного числа измерений, обладающее тем свойством, что грань каждого симплекса, входящего в наше множество симплексов, также входит в наше множество. Множества симплексов, удовлетворяющие этому условию, называются *симплициальными комплексами*; построенный нами симплициальный комплекс и есть, по определению, нерв системы множеств (1).

10*

З а м е ч а н и е. Симплексы нашего комплекса кроме примыканий по общим граням могут еще и пересекаться. Однако эти пересечения (а также *вырождения симплексов*, т. е. случаи, когда вершины какого-нибудь r-мерного симплекса оказываются лежащими в подпространстве размерности $< r$ основного пространства) могут быть легко избегнуты сколь угодно малым смещением вершин нашего комплекса, если только само пространство R имеет достаточно большое число измерений (если комплекс имеет число измерений n, т. е. если n есть наибольшее число измерений входящего в этот комплекс симплекса, то достаточно предположить, чтобы размерность пространства R была не менее, чем $2n + 1$; если пространство R имеет меньшую размерность, чем $2n+1$, но не меньшую, чем n, то сколь угодно малым сдвигом вершин комплекса можно достигнуть, чтобы ни один из симплексов не вырождался; но пересечения симплексов устранить, вообще говоря, нельзя). На фиг. 4a изображена система, состоящая из конечного числа квадратов, и ее нерв (число измерений его равно 2); система квадратов, изображенная на фиг. 4b, имеет в качестве нерва трехмерный комплекс.



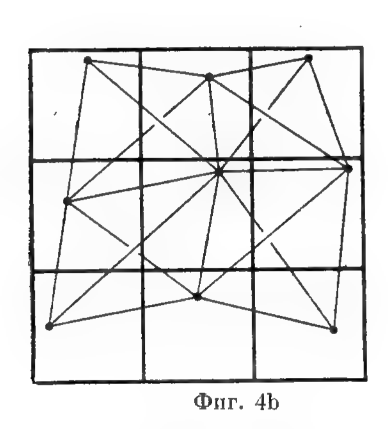
Фиг. 4а

Имеет место следующая не трудно доказываемая теорема:

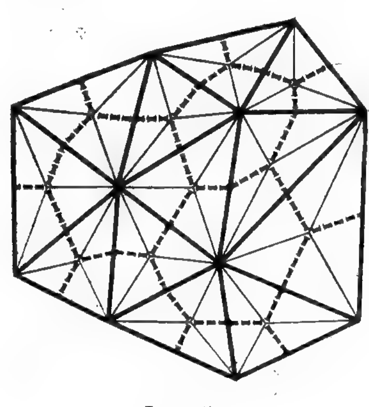
Пусть Π — *полиэдр*, n — *его число измерений и* K — *симплициальный комплекс, полученный, если разбить полиэдр* Π *на симплексы. Существует покрытие*[1] α *полиэдра* Π (*состоящее из замкнутых множеств и даже из полиэдров*), *имеющее своим нервом комплекс* K.

[1] Под покрытием какого-либо компакта Φ понимается конечная система α замкнутых подмножеств множества Φ, дающая в сумме множество Φ; покрытие α называется ε-покрытием, если все множества, являющиеся его элементами, имеют диаметр

Для получения покрытия α надо взять так называемое *барицентрическое подразделение* комплекса K (оно изображено в случае $n=2$ на фиг. 5 тонкими линиями) и рассмотреть соответствующие (обведенные на фиг. 5 жирным пунктиром) «барицентрические звезды» — они и являются элементами искомого покрытия α. Заметим, что, беря симплициальное раз-



Фиг. 4b



Фиг. 5

биение K достаточно мелким, можно при любом выборе положительного числа ε достигнуть, чтобы покрытие полиэдра Π, нервом которого является K, было ε-покрытием.

Пусть Φ произвольный компакт, лежащий в пространстве R и пусть дано отображение f, являющееся ε-сдвигом компакта Φ в полиэдр Π. Возьмем столь мелкое разбиение K полиэдра Π на симплексы, чтобы соответствующее «барицентрическое покрытие» нашего полиэдра (покрытие, имеющее нервом комплекс K) было ε-покрытием. Пусть B_1, B_2, ..., B_s суть элементы этого покрытия. Обозначим через Φ_i полный прообраз множества B_i при отображении f, т. е. замкнутое множество всех точек Φ,

---

$\leqslant \varepsilon$ (при этом диаметр какого-либо множества $\leqslant \varepsilon$, если расстояние между любыми двумя точками этого множества $\leqslant \varepsilon$); напоминаем, что кратностью покрытия называется наибольшее такое число $n+1$, что имеется $n+1$ множеств, являющихся элементами покрытия и имеющих непустое пересечение. Таким образом, кратность покрытия на 1 больше, чем число измерений нерва этого покрытия.

отображающихся посредством f в точки множества B_i. Легко видеть, что эти множества Φ_i образуют 3ε-покрытие компакта Φ, имеющее своим нервом тот же комплекс K. Таким образом, мы приходим к следующему заключению:

Всякий комплекс, являющийся достаточно мелким симплициальным разбиением полиэдра Π, *в который при некотором* ε*-сдвиге переходит данный компакт* Φ, *есть не что иное, как нерв некоторого* 3ε*-покрытия компакта* Φ.

Можно доказать и своего рода обратное предложение: *Если дано* ε*-покрытие* α *компакта* Φ, *то можно путем* ε*-сдвига отобразить* Φ *на полиэдр* Π, *составленный из всех или некоторых симплексов нерва данного* ε*-покрытия* (в очень широких предположениях, можно достигнуть того, чтобы полиэдр Π состоял из всех симплексов нерва α).

Таким образом, полиэдры, в которые данный компакт переходит путем малых сдвигов, оказываются поставленными в связь с нервами достаточно мелких покрытий этого компакта.

После всего сказанного открывается естественный путь к перенесению на произвольные компакты понятий, установленных для простейших геометрических фигур — полиэдров. Проследим этот путь на простейшем из этих понятий — на понятии размерности или числа измерений. Считая, что это понятие для полиэдров определено элементарным образом (как это делается, например, в аналитической или многомерной элементарной геометрии), мы определяем размерность произвольного компакта Φ как наименьшее неотрицательное целое число n, удовлетворяющее тому условию, что при любом $\varepsilon > 0$ компакт Φ можно помощью ε-сдвига преобразовать в n-мерный полиэдр.

Из связи, установленной нами между ε-сдвигами и ε-покрытиями, легко следует, что полученное нами, так сказать, из элементарно геометрических соображений общее понятие размерности может быть сформулировано и так:

Размерностью компакта Φ *называется наименьшее целое число* n, *обладающее тем свойством, что при всяком* $\varepsilon > 0$ *существует* ε*-покрытие компакта* Φ, *имеющее кратность* $n+1$.

Таким образом, мы приходим к определению размерности в той его форме, которая была дана одним из создателей теории размерности, —

П. С. Урысоном.[1] Из этого определения размерности легко следует, что два гомеоморфных между собою компакта имеют ту же размерность (другими словами, размерность обладает свойством топологической инвариантности).

При всей естественности и простоте общего понятия размерности его введение будет лишь тогда оправдано, если мы докажем, что всякий n-мерный в элементарно-геометрическом смысле полиэдр будет n-мерным и в смысле общего определения. А для этого на основании предыдущего достаточно доказать одно из следующих предложений, эквивалентных между собою в силу наших предыдущих рассуждений:

1° Для всякого n-мерного, в элементарном смысле этого слова, полиэдра можно найти такое $\varepsilon > 0$, что никаким ε-сдвигом этот полиэдр не может быть переведен в полиэдр меньшего числа измерений;

2° Для всякого n-мерного, в элементарном смысле этого слова, полиэдра можно найти такое $\varepsilon > 0$, что всякое ε-покрытие этого полиэдра имеет кратность $\geqslant n+1$.

Любое из этих предложений достаточно доказать для простейшего n-мерного полиэдра, именно для n-мерного симплекса T^n. Наметим, как это делается. Пусть дан n-мерный симплекс T^n. Возьмем $\varepsilon > 0$ столь малым, чтобы никакое множество, диаметра $< \varepsilon$, не могло иметь общих точек со всеми $(n-1)$-мерными гранями симплекса T^n. Отсюда, в частности, следует, что никакое множество диаметра $< \varepsilon$ не может одновременно содержать вершину e_i симплекса T^n и хотя бы одну точку грани T_i^{n-1} противоположной этой вершине. Пусть теперь

$$\Phi_0, \Phi_1, \ldots, \Phi_s \tag{2}$$

есть какое-нибудь ε-покрытие симплекса T^n. Предположим, что вершина e_i, $i = 0, 1, \ldots, n$, симплекса T^n содержится во множестве Φ_i. Если $s > 1$, то возьмем какое-нибудь Φ_j, $j > n$, и поступим следующим образом: отыщем грань T^{n-1} симплекса T^n, не содержащую ни одной точки мно-

[1] Павел Самуилович Урысон, один из наиболее блестящих по своим дарованиям советских математиков, родился в 1898 г. в Одессе, учился и работал в Московском университете; погиб от несчастного случая во время купанья в Бретани 17 августа 1924 г. Урысон принадлежит к числу крупнейших современных топологов; созданием теории размерности он проложил новые пути в развитии топологии.

жества Φ_j (такая грань в силу наших предположений существует); затем вычеркнем из системы (2) множество Φ, а множество Φ_i заменим через $\Phi_i+\Phi_j$, обозначив последнее множество снова через Φ_i. Вследствие этого число множеств (2) уменьшится на единицу, а кратность системы (2), как легко проверить, не повысится и не нарушится условие, чтобы ни одно из множеств (2) не содержало одновременно какую-нибудь вершину и точки противоположной этой вершине грани симплекса T^n. Повторяя этот процесс конечное число раз, мы получим, наконец, систему

$$\Phi_0, \Phi_1, \ldots, \Phi_n \tag{3}$$

из $n+1$ множества, являющуюся покрытием симплекса T^n, причем Φ_i содержит вершину e_i и не содержит ни одной точки грани T_i^{n-1}, противоположной вершине e_i. Отсюда легко выводится, что каждая $(n-1)$-мерная грань $|e_{i_0} \ldots e_{i_{n-1}}|$ симплекса T^n покрыта множествами $\Phi_{i_0}, \Phi_{i_1}, \ldots, \Phi_{i_{n-1}}$, соответствующим вершинам этой грани. Наконец, порядок системы (3) не больше, чем порядок системы (2). Следовательно, для того чтобы доказать, что порядок системы (2) не меньше $n+1$, нам достаточно доказать то же утверждение для системы (3), т. е. доказать, что имеется точка, принадлежащая всем множествам $\Phi_0, \Phi_1, \ldots, \Phi_n$ системы (3). Итак, все свелось к доказательству следующей теоремы:

Пусть n-мерный симплекс $T^n = |e_0\, e_1 \ldots e_n|$ покрыт замкнутыми множествами $\Phi_0, \Phi_1, \ldots, \Phi_n$, взаимно однозначно соответствующими его вершинам, так что вершина e_i содержится в Φ_i при любом $i = 0, 1, \ldots, n$ и, более того, каждая $(n-1)$-мерная грань $|e_{i_0} \ldots e_{i_{n-1}}|$ симплекса T^n покрыта множествами $\Phi_{i_0}, \Phi_{i_1}, \ldots, \Phi_{i_{n-1}}$. Тогда существует по крайней мере одна точка симплекса T^n, принадлежащая всем множествам Φ_0, $\Phi_1, \ldots, \Phi_n$.

Для доказательства этой теоремы достаточно показать, что на какие бы мелкие симплексы $t_1^n, t_2^n, \ldots, t_s^n$ мы ни разбили наш симплекс T^n, среди этих симплексов найдется по крайней мере один, имеющий общие точки с любым из множеств $\Phi_0, \Phi_1, \ldots, \Phi_n$. Покажем, что это последнее предложение легко сводится к ранее сформулированной нами лемме Шпернера. В самом деле, возьмем произвольное разбиение симплекса T^n на сколь угодно мелкие симплексы

$$t_1^n, t_2^n, \ldots, t_s^n.$$

Каждой вершине этого разбиения поставим в соответствие одно из содержащих эту вершину множеств $\Phi_0, \Phi_1, \ldots, \Phi_n$. Так как эти множества взаимно однозначно соответствуют вершинам основного симплекса T^n, то каждой вершине нашего разбиения поставлена в соответствие некоторая определенная вершина симплекса T^n, причем в силу предположений относительно системы множеств $\Phi_0, \Phi_1, \ldots, \Phi_n$ вершине разбиения, лежащей на какой-либо $(n-1)$-мерной грани симплекса T^n, поставлена в соответствие вершина именно этой грани. Мы находимся в условиях леммы Шпернера; следовательно, существует такой симплекс t^n нашего разбиения, вершинам которого поставлены в соответствие все вершины симплекса T^n; а это значит, что каждое из множеств Φ_i содержит некоторую вершину симплекса t^n, т. е. что симплекс t^n пересекается со всеми множествами $\Phi_0, \Phi_1, \ldots, \Phi_n$, что и требовалось доказать.

Итак, n-мерные симплексы, а значит и вообще n-мерные в элементарно геометрическом смысле полиэдры являются n-мерными и в смысле общего определения размерности. Из этого основного факта вытекают дальнейшие следствия. Прежде всего, так как два гомеоморфных компакта имеют всегда одинаковую размерность, а размерность симплекса равна его элементарному числу измерений, то два симплекса (и вообще два полиэдра) разного числа измерений не могут быть гомеоморфными между собою. А отсюда легко следует и знаменитая теорема Брауэра об *инвариантности числа измерений эвклидова пространства,* утверждающая невозможность при $m \neq n$ топологического отображения m-мерного эвклидова пространства (и вообще m-мерного многообразия) на n-мерное. При этом любопытно отметить, что конструктивная суть доказательства лежит в лемме Шпернера, которая при всей своей элементарности выражает действительно глубокий геометрический факт; все остальное представляет собою несколько упражнений на тему «замкнутые множества».

Доказательством равенства между размерностью полиэдра и его элементарно геометрическим числом измерений не ограничиваются приложения леммы Шпернера. Так, из леммы Шпернера легко вытекает инвариантность открытого множества m-мерного эвклидова пространства при топологическом отображении этого множества на какое-либо другое множество, также лежащее в m-мерном эвклидовом пространстве. Далее, простым следствием леммы Шпернера является и фундаментальная теорема Брауэра, утверждающая, что при всяком непрерывном отображе-

нии n-мерного выпуклого многогранника в себя всегда имеется хотя бы одна неподвижная точка (т. е. точка, переходящая при этом отображении в самое себя). Эта теорема Брауэра является первой в длинном ряду теорем о существовании неподвижных точек при непрерывных отображениях тех или иных фигур. Внимание топологов к теоремам этого рода объясняется не только собственно геометрическим интересом, но и в неменьшей мере многочисленными приложениями этих теорем (в особенности их бесконечномерных обобщений) к доказательствам разнообразнейших теорем существования в анализе.[1]

Однако вернемся к общему понятию размерности. Теорема о совпадении общей размерности полиэдра с его элементарным числом измерений, конечно, является необходимым условием для того, чтобы интересоваться дальнейшими свойствами общей размерности. Однако едва ли одно это условие достаточно для того, чтобы считать оправданным построение обширной теории, посвященной этому понятию; между тем теория размерности справедливо почитается и нашими и зарубежными учеными выдающимся достижением советской математики, занимающим в современной топологии одно из центральных мест. Дело в том, что основной интерес теории размерности заключается в том обилии новых конкретных геометрических фактов и неожиданных сопоставлений в области

---

[1] Чтобы понять, каким образом теоремы о существовании неподвижных точек при непрерывных отображениях могут привести к теоремам существования в анализе, рассмотрим простейшее дифференциальное уравнение $\frac{dy}{dx} = f(x,y)$ или эквивалентное ему интегральное уравнение

$$y(x) = \int_{x_0}^{x} f(t, y(t))dt. \qquad (4)$$

Правая часть уравнения ставит в соответствие каждой непрерывной функции $y = y(x)$ непрерывную функцию $Y(x) = \int_{x_0}^{x} f(t, y(t))\,dt$ и осуществляет этим путем непрерывное отображение *пространства непрерывных функций* в себя. Неподвижной точке этого отображения, очевидно, соответствует функция $y = y(x)$, являющаяся решением уравнения (4).

наиболее общих геометрических образований, к которым она приводит и которые обогащают геометрическое познание в такой же мере, в какой наше познание в области основных понятий анализа, каковы понятия функции, производной, интеграла, тригонометрического ряда и т. д., было обогащено возникновением и развитием теории функций вещественного переменного. Из многочисленных конкретных теорем, составляющих содержание теории размерности, я приведу лишь одну, а именно так называемую теорему о существенных отображениях, доказанную мною в 1930 г. и являющуюся основой так называемой гомологической теории размерности.

Возьмем какой-нибудь n-мерный компакт Φ, расположенный в m-мерном эвклидовом пространстве, и пусть положительное число ε настолько мало, что компакт Φ посредством ε-сдвига невозможно отобразить ни в какой полиэдр размерности $<n$. Выбрав такое ε, рассмотрим ε-сдвиг f_0 компакта Φ в какой-нибудь n-мерный полиэдр Π. Разобьем полиэдр Φ на симплексы диаметра $<\varepsilon$; полученный комплекс обозначим через K. Возьмем теперь какой-нибудь n-мерный симплекс T^n комплекса K, границу его обозначим через S^{n-1}; обозначим через A множество всех точек Φ, отображающихся посредством f_0 в точки симплекса T^n (при этом, как всегда, симплекс T^n считаем замкнутым, т. е. содержащим свою границу S^{n-1}). Обозначим, далее, через B множество точек Φ, отображающихся посредством f_0 в точки границы S^{n-1}, через C множество всех остальных точек Φ (т. е. множество всех тех точек Φ, которые посредством f_0 отображаются в точки полиэдра Π, не принадлежащие симплексу T^n). Назовем, наконец, *допустимой деформацией* отображения f_0 всякое такое непрерывное видоизменение f_t отображения f_0, зависящее от параметра t, $0 \leqslant t \leqslant 1$, при котором отображение в точках $B+C$ остается неизменным.[1]

---

[1] Подробнее: деформацией непрерывного отображения f_a компакта Φ на какое-нибудь множество (в нашем случае на полиэдр Π) называется семейство $\{f_t\}$ непрерывных отображений $f_t(x)$ множества Φ во множество Π, непрерывно зависящее от параметра t, $a \leqslant t \leqslant b$, в том смысле, что при достаточно малой разности $t' - t''$ точки $f_{t'}(x)$ и $f_{t''}(x)$ для любого x, принадлежащего Φ (т. е. равномерно по x) сколь угодно мало отстоят друг от друга. Считая параметр t временем, можно считать, что при изменении t от a до b отображение f_a, непрерывно изменяясь во времени («деформируясь»), переходит в отображение f_b.

Попытаемся допустимой деформацией отображения f_0 перевести это отображение в такое отображение f_1, при котором хотя бы одна точка симплекса T^n перестала быть образом какой-либо точки компакта Ф. Если такое «освобождение» хотя бы одной точки o симплекса T^n возможно, то, подвергая лежащую в T часть множества f_1 (Ф) центральной проекции из точки o, мы «сгоним» всю эту часть множества Ф на границу симплекса T, т. е. получим дальнейшую допустимую деформацию f_t, $1 \leqslant t \leqslant 2$, отображения f_1, в результате которой образ множества Ф не будет содержать ни одной внутренней точки симплекса T^n. При этом, если x — точка множества A (т. е. если $f_0(x)$ принадлежит T^n), то при любом t, $0 \leqslant t \leqslant 2$, точка $f_t(x)$ также лежит в симплексе T^n (вся деформация, переводящая $f_0(x)$ в $f_2(x)$ протекает в пределах симплекса T^n); если же x — точка множества $C = Ф—(A+B)$, то при любом t точка $f_t(x)$ совпадает с $f_0(x)$. Так как диаметр любого симплекса T^n комплекса K меньше ε, то из сказанного следует, что отображение f_2 есть 2ε-сдвиг, поэтому образ компакта Ф при отображении f_2 не может быть полиэдром размерности $\leqslant n—1$. А отсюда следует, что среди n-мерных симплексов T^n комплекса K непременно найдутся такие, которые при любой допустимой деформации отображения f_0 будут сплошь покрыты образами точек множества Ф.

Теперь естественно ввести следующее определение. Пусть дано непрерывное отображение f_0 компакта Ф на n-мерный симплекс T^n, границу которого мы снова обозначаем через S^{n-1}. Обозначим через B множество всех точек Ф, которые при отображении f_0 переходят в точки границы S^{n-1} симплекса T^n, и назовем допустимой всякую деформацию f_t, $0 \leqslant t \leqslant 1$, отображения f_0, при которой в точках множества B отображение остается неизменным. Отображение f_0 компакта Ф на симплекс T^n называется *существенным*, если при всякой допустимой деформации этого отображения все точки симплекса T^n продолжают быть образами точек компакта Ф.

Из приведенных выше рассуждений следует, что при рассмотренном отображении f_0 компакта Ф на полиэдр П найдется такой симплекс T^n этого полиэдра, что на этот симплекс существенно отображается некоторая часть (а именно множество A) компакта Ф. Отсюда легко следует, что и весь компакт Ф можно существенно отобразить на симплекс T^n. Итак, *каждый n-мерный компакт* Ф *можно существенно отобразить на*

n-мерный симплекс. С другой стороны, можно доказать, что всякое непрерывное отображение любого n-мерного компакта на симплекс размерности $> n+1$ всегда является несущественным.[1]

Итак, *возможность существенных отображений на симплексы того или иного числа вполне характеризует размерность: размерностью данного компакта Ф является наибольшее целое число n, обладающее тем свойством, что компакт Ф можно существенно отобразить на n-мерный симплекс.* Эта теорема (так же, как возможность определить размерность помощью ε-сдвигов) выясняет наглядно геометрическое содержание понятия размерности и подтверждает слова Гильберта о том, что даже в самых общих построениях современной топологии сохраняется связь с непосредственным пространственным воззрением.

Кстати, несколько слов о степени общности наших построений. Чтобы не затруднять читателя, мы ограничили наши рассмотрения компактами, лежащими в m-мерном эвклидовом пространстве, рассматривая множества, расположенные в гильбертовом пространстве, лишь когда речь шла о самых элементарных топологических понятиях. В действительности, если ограничиваться метрическими пространствами, то введенные нами ограничения являются лишь словесными: в силу одной из известнейших теорем Урысона, *всякое* компактное метрическое пространство гомеоморфно некоторому множеству, лежащему в гильбертовом пространстве. Далее, всякое компактное метрическое пространство Ф *конечной размерности* гомеоморфно множеству, лежащему в эвклидовом пространстве достаточно большого числа измерений (именно, если размерность Ф равна n, то Ф гомеоморфно некоторому множеству, лежащему в $(2n+1)$-мерном эвклидовом пространстве).

Эта фундаментальная теорема была впервые доказана Небелингом при широком использовании того самого метода ε-сдвигов, которому в значительной степени посвящен настоящий параграф этой статьи. Дру-

---

[1] Например, так называемые «кривые Пеано» сводятся к тому, чтобы задать две непрерывные функции $x = f(t)$, $y = \varphi(t)$ так, что при t, пробегающем отрезок $0 \leqslant t \leqslant 1$, точка (x, y) пробегает множество всех точек некоторого треугольника. Другими словами, задать кривую Пеано значит задать некоторое непрерывное отображение отрезка на треугольник или квадрат; *все такие отображения несущественны*: это легко следует из возможности с произвольной точностью приблизить непрерывную функцию «кусочно-линейными» функциями.

гими авторами (в том числе Л. С. Понтрягиным) были предложены другие доказательства теоремы Небелинга, но все эти доказательства существенно опираются на тот же метод ε-сдвигов. Это вполне естественно: из совершенно элементарных соображений аналитической геометрии вытекает, что всякий n-мерный полиэдр, заданный абстрактно (комбинаторной схемой его разбиения на симплексы или как множество, лежащее, скажем, в гильбертовом пространстве), гомеоморфен замкнутому множеству $(2n+1)$-мерного пространства. Речь, стало быть, идет о том, чтобы это элементарное и давно известное свойство полиэдров распространить на n-мерные компакты, т. е. *на множества, которые сколь угодно малыми сдвигами могут быть преобразованы в n-мерные полиэдры*; рассмотрение полиэдров, аппроксимирующих данный компакт в изложенном выше смысле, занимает поэтому центральное место во всех доказательствах теоремы Небелинга. Важность этой теоремы заключается в том, что она позволяет среди всех метрических и даже среди всех топологических пространств[1] полностью охарактеризовать те, которые гомеоморфны множествам, лежащим в эвклидовых пространствах: *для этого кроме так называемой нормальности и наличия счетной базы* (гарантирующих в силу результатов Урысона возможность топологически отображения в гильбертово пространство) *необходимо и достаточно, чтобы пространство имело конечную размерность.*

§ 4

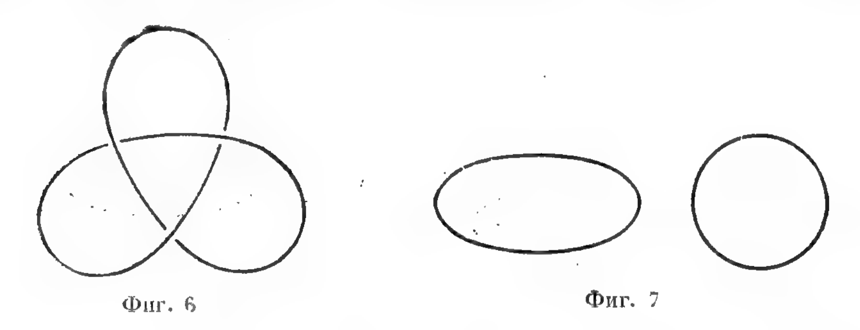
ПОНЯТИЕ ЗАЦЕПЛЕНИЯ. ЕЩЕ ОБ m-МЕРНОЙ ТЕОРЕМЕ ЖОРДАНА

Два множества, лежащих в данном m-мерном эвклидовом пространстве, как мы знаем, называются гомеоморфными (или, с греческого, *имеющими сходную форму*), если одно из них можно топологически отобразить

---

[1] Определения топологического пространства, нормальности, базы, равно как доказательство упоминаемой здесь теоремы Урысона, можно найти в главе 6 книги Хаусдорфа «Теория множеств». Определение размерности для любого нормального пространства дано в главе 1 моей книги «Комбинаторная топология» и в моей статье «Понятие пространства в топологии» (Усп. матем. наук, 1947, т. 2, вып. 1).

на другое. Два множества называются изотопными (по отношению к тому пространству, в котором они оба лежат) или имеющими одинаковое расположение в этом пространстве, если все пространство можно топологически отобразить на себя так, чтобы при этом отображении одно из наших множеств отобразилось на другое. В начале предыдущего параграфа была упомянута одна из теорем Антуана, в силу которой, как мы теперь можем сказать, всякая простая замкнутая линия, лежащая в плоскости,



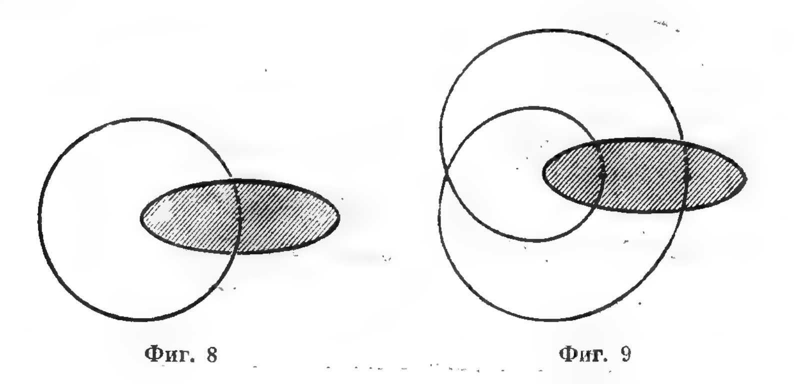
Фиг. 6 Фиг. 7

изотопна по отношению к этой плоскости обыкновенной окружности. В трехмерном пространстве простейшим примером простой замкнутой линии неизотопной окружности является пример заузленной замкнутой линии (фиг. 6).

Столь же простые примеры гомеоморфных, но не изотопных фигур в трехмерном пространстве можно получить, если взять с одной стороны две незацепленные замкнутые кривые (фиг. 7), а с другой — две зацепленные (фиг. 8 или фиг. 9). При этом две окружности, изображенные на фиг. 8, имеют коэффициент зацепления, равный ± 1 (знак зависит от выбранной ориентации наших окружностей, т. е. от выбранных на них направлений обхода, в соединении с выбранной ориентацией всего пространства), а две замкнутые кривые, изображенные на фиг. 9, имеют коэффициент зацепления ± 2. Наконец, две незацепленные окружности имеют коэффициент зацепления, равный нулю.

Явление зацепления двух замкнутых кривых, расположенных в трехмерном пространстве, заключается в том, что каждая из двух данных кривых пересекает всякую поверхность (всякую «пленку»), натянутую на вторую кривую. При этом можно точкам пересечения поверхности

с кривой некоторым определенным образом приписывать знак + или — (в этом правиле знаков учитываются как направление обхода кривых, так и ориентация куска поверхности и, наконец, ориентация всего пространства) и тогда говорить об алгебраическом числе пересечений данной поверхности с данной кривой, понимая под этим числом разность между числом положительных и числом отрицательных точек пересечения.



Фиг. 8 Фиг. 9

Оказывается, если в трехмерном пространстве даны две замкнутые кривые без общих точек, то алгебраическое число пересечений одной из данных двух кривых с поверхностью, натянутой на другую, не зависит от специального выбора этой поверхности и выражает некоторое свойство расположения обеих данных кривых в пространстве.

Определенное таким образом число и называется коэффициентом зацепления данных кривых. Аналитически он выражается известным интегралом Гаусса.

Понятие зацепления обобщается и на случай замкнутых многообразий, расположенных в m-мерном пространстве.

Считая пару точек *нульмерным многообразием* («нульмерной сферой»), можно рассматривать разбиение плоскости замкнутой линией или разбиение пространства замкнутой поверхностью как частный случай зацепления: в самом деле, если, например, в трехмерном пространстве дана замкнутая поверхность Ф, а из двух точек a и b одна лежит внутри поверхности, а другая — во внешней области к этой поверхности, то

каждый непрерывный путь, соединяющий точки a и b («каждая линия, ограниченная парой точек a, b», или каждая «одномерная пленка, натянутая на эту пару точек»), пересекается с поверхностью Ф. Точно так же внутренняя область к нашей поверхности содержит одну из заданных двух точек, т. е. пересекается с нашим нульмерным многообразием; мы видим, что явление зацепления поверхности с парой точек имеет в трехмерном пространстве полную силу.

Идея рассматривать разбиение пространства как частный случай зацепления впервые была высказана Лебегом в 1911 г. Эта идея оказалась чрезвычайно плодотворной: с большой силой примененная Александером к новым, гораздо более широким вопросам, она стала основой так называемых топологических теорем двойственности, простейшей из которых и является теорема Жордана. Сам Лебег применил свою точку зрения на разбиение пространства как на частный случай зацепления к доказательству n-мерной теоремы Жордана, точнее, утверждения, что n-мерное многообразие разбивает $(n+1)$-мерное пространство по крайней мере на две области. Доказывается более общее предложение, именно:

Ко всякому p-мерному замкнутому ориентируемому многообразию (вообще, ко всякому p-мерному циклу),[1] лежащему в m-мерном эвклидовом пространстве, можно найти зацепленное с ним $(m-p-1)$-мерное многообразие.

Наметим общий ход доказательства в случае $m=3$.

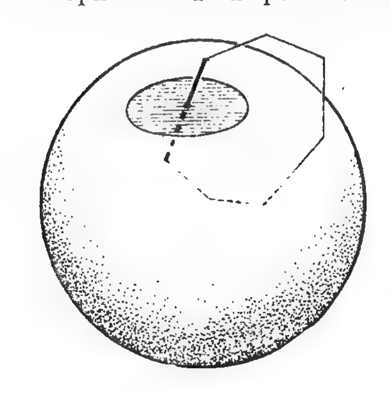
При $p=0$ теорема очевидна: нульмерное многообразие есть, по определению, пара точек; заключая одну из этих точек внутрь шара, столь малого, что вторая точка оказывается вне шара, и беря поверхность этого шара, получим двумерное многообразие, зацепленное с данным нульмерным.

Пусть теперь $p=1$; одномерное многообразие есть простая замкнутая линия C; требуется построить к ней зацепленную с ней простую замкнутую линию C'. Для этого берем на C' пару точек a, b и обозначаем через Λ одну из тех двух дуг, на которые пара точек a, b разбивает линию C. Строим теперь зацепленную с парой точек a, b сферическую поверхность S. Она пересекает каждую из дуг, соединяющих точки a и b,

[1] См. § 3.

в частности и дугу Λ. Пересечение этой дуги с поверхностью S образует лежащее на этой поверхности замкнутое множество. Это замкнутое множество можно покрыть конечным числом треугольников, взятых из некоторой достаточно мелкой триангуляции поверхности S. Эти треугольники образуют на поверхности одну или несколько связных областей, которые можно предположить ограниченными несколькими замкнутыми линиями. По крайней мере одна из этих замкнутых линий и оказывается зацепленной с замкнутой линией C.

Наконец, в случае $p=2$ имеем в трехмерном пространстве замкнутую поверхность Ф. Берем на ней какую-нибудь точку a и малую окрестность этой точки, гомеоморфную кругу и ограниченную простой замкнутой линией C. Строим простую замкнутую линию C', зацепленную с C. Она пересекает всякую пленку, натянутую на C, значит, в частности, и тот ограниченный линией C кусок поверхности S, который является окрестностью точки a. Этот кусок поверхности обозначим через E. Пересечение E с C' дает на C' замкнутое множество, которое можно покрыть конечным числом дуг, лежащих на C' и попарно не пересекающихся. Каждая из этих дуг ограничена парой точек и по крайней мере одна из этих пар зацеплена с поверхностью Ф.



Фиг. 10

На фиг. 10 пересечение E с C' состоит из одной точки, и пара точек, ограничивающих дугу линии, содержащую эту точку пересечения, зацеплена с поверхностью S.

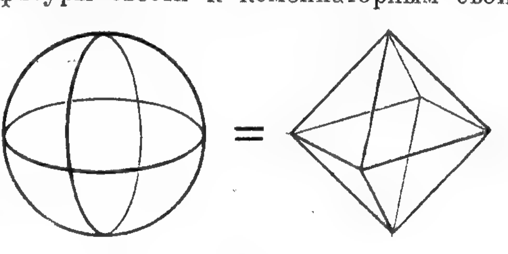
Наше рассуждение, передающее построение Лебега, есть, очевидно, не что иное, как индукция от p к $p+1$; этой же индукцией теорема доказывается и в случае любого n.

Теперь мы уже совсем подошли к знаменитому закону двойственности Александера. Но прежде чем дать хотя бы формулировку закона двойственности Александера, нам необходимо, наконец, аккуратно ввести основные вспомогательные понятия так называемой комбинаторной топологии.

§ 5

ОСНОВНЫЕ ПОНЯТИЯ КОМБИНАТОРНОЙ ТОПОЛОГИИ

Комбинаторная топология в широком смысле слова может быть определена как исследование тех или иных геометрических фигур самой различной степени общности (начиная от простейших полиэдров до наиболее общих топологических пространств) путем *рассмотрения их покрытий*, т. е. представлений данной фигуры в виде суммы конечного числа достаточно «малых» ее кусков. Таким образом удается значительную часть топологических свойств данной фигуры свести к комбинаторным свойствам той или иной конечной схемы, характеризующей структуру данного покрытия (т. е. имеющиеся в нем отношения взаимного соседства — прилегания или пересечения — отдельных его элементов). Например, комбинаторная топология поверхностей пользуется в качестве основного средства изучения так называемыми триангуляциями поверхностей, т. е разбиениями их на конечное число, вообще говоря, криволинейных треугольников (см. например, фиг. 11, на которой изображена одна из триангуляций сферы). Криволинейность треугольников, являющихся элементами триангуляции, при этом несущественна, так как всякая поверхность, допускающая триангуляцию, гомеоморфна некоторой полиэдральной поверхности, могущей быть разбитой уже на обыкновенные прямолинейные треугольники.



Фиг. 11

В многомерной топологии аналогичным образом исследуются кривые полиэдры (т. е. топологические образы обыкновенных многомерных полиэдров), в первую очередь многообразия, путем рассмотрения *комплексов*, получающихся при разбиении полиэдров на симплексы. Наконец, в случае более общих топологических пространств и прежде всего компактов (т. е. компактных метрических пространств) приходится рассматривать *произвольные* покрытия (так как разбить компакт на симплексы, вообще говоря, невозможно); но здесь на помощь приходит понятие нерва,

позволяющее свести дело опять-таки к рассмотрению некоторых симплициальных комплексов. В случае полиэдра для определения всех его топологических свойств принципиально достаточным является знание какого-нибудь хотя бы одного единственного комплекса, являющегося симплициальным разбиением данного полиэдра (между прочим, как мы видели, этот комплекс всегда является нервом некоторого вполне определенного покрытия данного полиэдра — именно покрытия барицентрического). В более общих случаях совокупность топологических свойств данного пространства (например, компакта) не может быть извлечена из знания нерва N_α одного какого-нибудь покрытия α: приходится рассматривать системы бесконечного числа покрытий; при этом, если покрытие β является подразделением покрытия α (т. е. если элементы покрытия α являются суммами содержащихся в них элементов покрытия β), то, ставя в соответствие каждому элементу покрытия β содержащий его элемент покрытия α, мы получим отображение не только вершин комплекса N_β на вершины комплекса N_α, но и всего комплекса N_β на комплекс N_α (т. е. каждому симплексу N_β соответствует симплекс комплекса N_α).

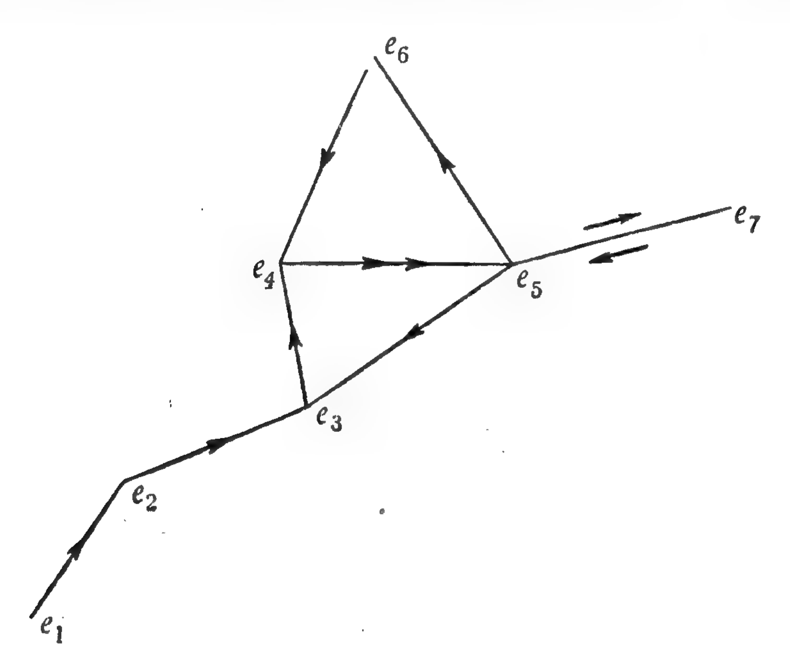
Эти отображения нервов более мелких покрытий на нервы более крупных называются *проекциями*. Оказывается, если задать совершенно отвлеченно совокупность нервов всех покрытий данного пространства и проекции этих нервов, то получающееся образование (так называемый «проекционный спектр») вполне определяет то пространство, из рассмотрений покрытий которого данный проекционный спектр получен. Таким образом, посредством проекционных спектров изучение даже чрезвычайно общих топологических пространств сводится к изучению симплициальных комплексов и их так называемых симплициальных отображений.

Итак, центральным вспомогательным понятием топологии оказывается понятие комплекса. В изучении комплексов можно пользоваться алгебраическими методами, с которыми мы сейчас и познакомимся вкратце.

В основе этих методов лежит понятие *цепи* (данного числа измерений r). Рассмотрим сначала случай $r=1$.

Возьмем какую-нибудь направленную ломаную линию, в которой могут быть и самопересечения, и отрезки, проходимые по нескольку раз,

но которая вся проходится нами в одном определенном направлении (см., например, на фиг. 12 ломаную $x^1 = e_1\, e_2\, e_3\, e_4\, e_5\, e_6\, e_4\, e_5\, e_7\, e_5\, e_3$). Каждый из отрезков, составляющих нашу ломаную, *взятый с определенным,*



Фиг. 12

произвольным направлением, которое вовсе не обязано совпадать с направлением нашего движения, обозначим как-нибудь: например, для случая, изображенного на нашем чертеже, положим

$$(e_1e_2)=t_1,\quad (e_2e_3)=t_2,\quad (e_3e_4)=t_3,\quad (e_4e_5)=t_4,$$
$$(e_5e_6)=t_5,\quad (e_4e_6)=t_6,\quad (e_3e_5)=t_7,\quad (e_5e_7)=t_8.$$

Тогда всю нашу ломаную $x = e_1\, e_2\, e_3\, e_4\, e_5\, e_6\, e_4\, e_5\, e_7\, e_5\, e_3$ можно записать в виде

$$x^1=t_1+t_2+t_3+t_4+t_5-t_6+t_4+t_8-t_8-t_7,$$

или

$$x^1=t_1+t_2+t_3+2t_4+t_5-t_6-t_7,$$

или, наконец,

$$x^1 = \sum_1^8 c_i\, t_i,$$

где коэффициент c_i в *линейной форме* $c_i\, t_i$ получается так: если по ходу нашего пути по ломаной направленный отрезок t_i проходится p_i раз в своем направлении, а q_i раз — в противоположном, то $c_i = p_i - q_i$. Итак, рассмотрение путей вдоль одномерных элементов (отрезков) данного комплекса K естественно приводит к рассмотрению линейных форм вида $x^1 = \Sigma c_i t_i^1$, где t_i^1 суть как-нибудь ориентированные одномерные симплексы комплекса K, а c_i — целые числа. Такие линейные формы и называются одномерными цепями комплекса K.

Точно так же линейные формы $x^2 = \Sigma c_i t_i^2$, где t_i^2 суть как-нибудь ориентированные (снабженные направлением обхода) двумерные симплексы (треугольники) комплекса K, а c_i — целые числа, называются двумерными цепями комплекса K.

Легко дать и общее определение. Прежде всего, r-мерным ориентированным симплексом называется симплекс, рассматриваемый вместе с некоторым определенным порядком его вершин; при этом два различных порядка вершин, переходящих друг в друга путем четной перестановки, определяют одну и ту же ориентацию. Если данная ориентация r-мерного симплекса $|e_0 e \ldots e_r|$ определяется, скажем, порядком e_0, e_1, $\ldots$, e_r его вершин, то ориентированный таким образом симплекс записывается в виде $t^2 = (e_0 e_1 \ldots e_r)$. Тогда порядок вершин $e_1\, e_0\, e_2 \ldots e_r$ определит противоположно ориентированный симплекс $-t^r = (e_1 e_0 e_2 \ldots e_r)$, а, например, порядок вершин $(e_1\, e_2\, e_0\, e_3 \ldots e_r)$ определит снова тот же ориентированный симплекс t^r. Симплекс, рассматриваемый без ориентации, будем обозначать через $|t^r|$ или $|e_0\, e_1 \ldots e_r|$.

Линейная форма вида $x^r = \Sigma\, c_i\, t_i^r$, где t_i^r суть ориентированные r-мерные симплексы данного комплекса K, а коэффициенты c_i суть целые числа, называется r-мерной целочисленной цепью комплекса K. Нульмерной цепью называем выражение вида $x^0 = \Sigma\, c_i\, t_i^0$, где t_i^0 суть вершины комплекса K, а c_i, как всегда, целые числа. Целочисленные r-мерные цепи комплекса K можно складывать, как обыкновенные линейные формы; среди них имеется «нулевая цепь», т. е. форма, в которой все коэффициенты суть нуль, наконец, к каждой цепи $x^r = \Sigma\, c_i\, t_i^r$ существует противо-

положная ей цепь $-x^r = \Sigma - c_i t_i^r$. Словом, r-мерные цепи образуют коммутативную группу,[1] являющуюся прямой суммой стольких бесконечных циклических групп, как велико число r-мерных симплексов данного комплекса K. Эта группа обозначается через $L^2(K)$.

Основное значение в комбинаторной топологии имеет операция образования границы цепи. В ее основе лежит совершенно элементарная геометрическая конструкция, которую мы сейчас и изложим. Естественно за границу ориентированного треугольника $(e_0\, e_1\, e_2)$ считать одномерную цепь $(e_0\, e_1) + (e_1\, e_2) + (e_2\, e_0)$. Точно так же граница ориентированного r-мерного симплекса

$$t^r = (e_0 e_1 \ldots e_r)$$

определяется как $(r-1)$-мерная цепь

$$\Delta t^r = \sum (-1)\, t_i^{r-1}, \quad \text{где} \quad t_i^{r-1} = (e_0 \ldots e_{i-1} e_{i+1} \ldots e_r).$$

Для одномерного симплекса $t^1 = (e_0\, e_1)$ получаем, в частности, $\Delta t^1 = e_1 - e_0$. После того как определена граница r-мерного ориентированного симплекса, мы определяем границу любой r-мерной цепи $x^r = \Sigma\, c_i\, t_i^r$ комплекса K по формуле

$$\Delta x^r = \sum c_i\, \Delta t_i^r; \tag{1}$$

помня, что каждое Δt_i^r есть $(r-1)$-мерная цепь, видим, что правая часть равенства (1) также есть $(r-1)$-мерная цепь. Итак, каждой r-мерной цепи x^r соответствует некоторая $(r-1)$-мерная цепь Δx^r. Этим определено гомоморфное отображение Δ группы $L^r(K)$ на некоторую подгруппу $H^{r-1}(K)$ группы $L^{r-1}(K)$. Элементы группы $H^{r-1}(K)$ — это те $(r-1)$-мерные цепи, которые являются границами некоторых r-мерных цепей комплекса K; эти $(r-1)$-мерные цепи называются цепями, *гомологичными нулю* в K. При гомоморфизме Δ на нулевой элемент группы $L^{r-1}(K)$ отображаются те цепи x^r, граница которых равна нулю. Эти цепи называются r-мерными *циклами* комплекса K; они образуют подгруппу $Z^r(K)$ группы $L^r(K)$.

Гомоморфизм Δ обладает тем замечательным свойством, что $\Delta\Delta x^r = 0$ для всякой цепи x^r; другими словами, граница каждой цепи есть цикл.

[1] Все необходимые для понимания этой статьи сведения по теории групп можно с избытком найти в главе I книги Л. С. Понтрягина «Непрерывные группы» (М., 1938). При этом во всей этой статье мы под группой понимаем всегда коммутативную группу.

Для доказательства этого утверждения достаточно показать, что граница всякого ориентированного симплекса есть цикл (т. е. что $\Delta\Delta t^r=0$); но в этом легко убедиться непосредственным вычислением. Очевидно, утверждение, что граница всякой цепи есть цикл, может быть сформулировано и так: группа $H^r(K)$ есть подгруппа группы $Z^r(K)$.

Рассмотрим теперь фактор-группу группы $Z^r(K)$ по ее подгруппе $H^r(K)$; ее элементами являются классы гомологичных между собою циклов (два цикла z_1^r и z_2^r называются гомологичными между собою, если их разность гомологична нулю). Фактор-группа $Z^r(K)/H^r(K)$ называется r-мерной *группой Бетти* комплекса K и обозначается через $\Delta^r(K)$. Эти группы определены (как и все наши только что сделанные построения) для любого комплекса K. При этом надо иметь в виду, что существует бесконечно много различных комплексов, являющихся симплициальными разбиениями одного и того же полиэдра П; эти комплексы будут обладать совершенно различными группами $L^r(K)$, $Z^r(K)$, $H^r(K)$; однако группы Бетти данного числа измерений у всех этих комплексов будут изоморфны между собою: *группы Бетти какого-нибудь симплициального разбиения полиэдра* П *являются топологическими инвариантами самого полиэдра* и при этом — можно не колеблясь это сказать,— вместе с размерностью *самыми важными топологическими инвариантами.*

Группы Бетти содержат в себе и некоторые числовые топологические инварианты полиэдра П, которые и были открыты основателем комбинаторной топологии Пуанкаре. Прежде всего, так как число симплексов комплекса K, являющегося симплициальным разбиением данного полиэдра П, — конечно, то группы $L^r(K)$, а следовательно, все подгруппы и фактор-группы этих групп суть группы с конечным числом образующих. Структура таких групп[1] может считаться полностью изученной. Всякая группа с конечным числом образующих имеет прежде всего конечный *ранг* p (ранг p определяется условием, что в данной группе можно найти линейно независимую систему, состоящую из p элементов, тогда как всякие $p+1$ элементов группы линейно зависимы). Далее, всякая группа с конечным числом образующих ранга p есть прямая сумма конечной подгруппы (состоящей из всех элементов данной группы, имеющих конечный порядок) и *свободной* группы ранга p (т. е. прямой суммы p

---

[1] См. подстрочное примечание на стр. 167.

бесконечных циклических групп). Применяя эти понятия к группе $\Delta^r(K)$, получаем прежде всего *r-мерное число Бетти* $p^r(K)$ *комплекса* K — это ранг группы $\Delta^r(K)$. Далее, подгруппа, состоящая из всех элементов конечного порядка группы $\Delta^r(K)$, есть конечная группа, называемая *r*-мерной группой *кручения* данного комплекса (соответственно полиэдра П). Ее элементы суть те классы гомологичных между собою циклов z^r, которые обладают следующим свойством: цикл z^r не гомологичен нулю в K, однако существует такое целое число $c > 1$, что цикл cz^r гомологичен нулю в K.

Рассмотрим несколько примеров. Если K — замкнутый полигон без кратных точек («окружность»), то группа $\Delta^1(K)$ есть бесконечная циклическая группа. Если K — триангуляция полиэдра, гомеоморфного сфере (например, комплекс, состоящий из всех граней тетраэдра), то группа $\Delta^1 K$ есть нульгруппа (состоит из одного нулевого элемента), группа $\Delta^2 K$ — бесконечная циклическая. Группа $\Delta^0 K$ для всякого связного комплекса есть бесконечная циклическая (вообще если комплекс распадается на s «связных кусков» (компонент), то его нульмерное число Бетти равно s).

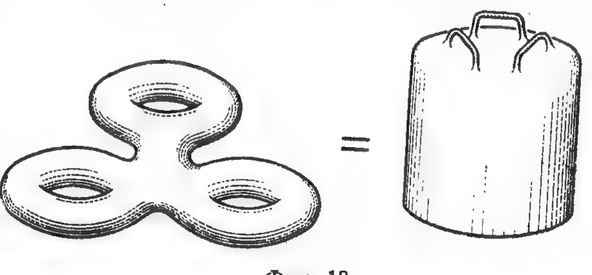
Итак, для любой триангуляции сферы нульмерное и двумерное числа Бетти равны 1, а одномерное число Бетти равно нулю. Для тора имеем: одномерное число Бетти равно 2, а нульмерное и двумерное равны 1.

Ни один комплекс не имеет отличной от нуля нульмерной группы кручения; кроме того, n-мерная группа кручения n-мерного комплекса также есть нульгруппа (это происходит от того, что для n-мерного комплекса группа $H^n(K)$ есть, очевидно, нульгруппа (в n-мерном комплексе K никакой отличный от нуля n-мерный цикл не может быть гомологичен нулю, так как ему нечего в этом комплексе ограничивать); поэтому группа $\Delta^n(K)$ совпадает с группой $Z^n(K)$, а эта последняя группа как подгруппа свободной группы $L^n(K)$ сама есть свободная группа и, значит, не имеет отличных от нуля элементов конечного порядка.

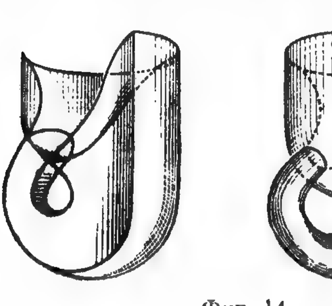
В силу изложенного, у поверхностей если и существуют отличные от нуля группы кручения, то только одномерные. Но у замкнутых поверхностей, без самопересечений помещающихся в трехмерном пространстве [таковы тор, а также поверхности рода p («сфера с p ручками» — на фиг. 13 изображен случай $p=3$)] группы кручения всегда суть нульгруппы. При этом одномерное число Бетти поверхности рода p равно $2p$. Так называемые неориентируемые замкнутые поверхности (например,

поверхность Клейна, изображенная на фиг. 14), будучи помещены в трехмерное пространство, всегда дают там самопересечения. Для неориентируемых замкнутых поверхностей одномерная группа кручения есть группа второго порядка.

Наши рассуждения могут быть обобщены в нескольких различных направлениях, которые все существенны и понадобятся нам в дальнейшем. Во-первых, наряду с конечными симплициальными комплексами можно рассматривать бесконечные комплексы. Бесконечное множество K, элементы которого суть симплексы различного числа измерений, называется бесконечным комплексом, если любая грань симплекса, являющегося элементом множества K, есть элемент множества K и если во множестве K не существует элемента, являющегося гранью бесконечного числа симплексов — элементов K (это последнее условие, называемое условием «локальной конечности», достаточно потребовать лишь в применении к нульмерным элементам, т. е. вершинам комплекса K; тогда оно само собою окажется выполненным и для элементов любого числа измерений). Для бесконечного комплекса K группы $L^r(K)$, $Z^r(K)$, $H^r(K)$, $\Delta^r(K)$ определяются совершенно так же, как для конечного комплекса (при этом цепи комплекса K определяются как *конечные* линейные формы $\Sigma c_i t_i^r$, т. е. требуется, чтобы в выражении $\Sigma c_i t_i^r$ лишь конечное число коэффициентов было отлично от нуля). Важность рассмотрения бесконечных комплексов вытекает хотя бы из того, что всякое открытое множество m-мерного эвклидова пространства R^m может быть разбито на симплексы так, что получится бесконечный комплекс. При этом, какое бы



Фиг. 13



Фиг. 14

разбиение данного открытого множества Г мы ни взяли, группы Бетти всегда с точностью до изоморфизма будут одни и те же; более того, они не будут меняться и при топологических отображениях открытого множества Г, следовательно являются топологическими инвариантами Г. Поэтому можно говорить о группах Бетти открытого множества Г пространства R^m, понимая под ними группы Бетти какого-нибудь (всё равно какого) симплициального разбиения этого множества.

Во-вторых, можно определить группы Бетти для любого компакта и даже для гораздо более общих топологических пространств. Для этого рассматривается система всех покрытий α данного пространства Ф и нервы N_α этих покрытий. Если покрытие β является подразделением покрытия α, то, как уже упоминалось, определено так называемое симплициальное отображение π_α^β комплекса N_β на комплекс N_α (проекция). При этом отображении группа $\Delta^r(N_\beta)$ гомоморфно отображается в группу $\Delta^r(N_\alpha)$ так, что получается система групп, связанных гомоморфизмами[1]. Такая система групп по определенным правилам порождает некоторую предельную группу, которая и называется группой Бетти Δ^r(Ф) пространства Ф. Этот способ определения групп Бетти применим к очень широкому классу топологических пространств и с теоретической точки зрения заслуживает предпочтения перед всеми другими. Однако если ограничиваться лишь компактами и рассматривать их как множества, лежащие в эвклидовом или в гильбертовом пространстве R, то удобнее следующий способ. В пространстве R, в котором лежит компакт Ф, рассматриваются всевозможные конечные системы ориентированных симплексов

$$t_1^r, \ldots, t_s^r$$

данного числа измерений r с *вершинами, являющимися точками* Ф; линейные формы вида $x^r = \Sigma\, c t_i^r$ называются цепями компакта Ф; цепь $x^r = \Sigma\, c_i t_i^r$ называется ε-цепью компакта Ф, если диаметры всех входящих в нее симплексов меньше данного $\varepsilon > 0$. Очевидно, что граница Δx^r всякой ε-цепи тоже есть ε-цепь (даже ε-цикл: ε-цепь называется ε-циклом, если ее граница равна нулю). Пусть теперь δ и ε — два положительных числа. Мы говорим, что данный δ-цикл z^r компакта Ф является ε-гомологичным нулю в этом компакте, если можно подобрать ε-цепь этого

[1] Которые впервые установил выдающийся советский тополог и алгебраист Л. С. Понтрягин.

компакта, имеющую своей границей данный δ-цикл; само собою разумеется, что два δ-цикла называются ε-гомологичными между собою в Φ, если их разность ε-гомологична нулю в Φ. Последовательность

$$z^r = (z_1^r, z_2^r, \ldots, z_k^r, \ldots), \tag{2}$$

где z_k^r есть δ_k-цикл компакта Φ и $\delta_k \to 0$, называется *истинным циклом* компакта Φ, если, каково бы ни было $\varepsilon > 0$, существует такое k, что для любых $p > k$, $q > 0$ циклы z_p^r и z_q^r между собою ε-гомологичны в Φ. Наконец, истинный цикл (4) называется гомологичным нулю в Φ, если для любого $\varepsilon > 0$ найдется такое n_ε, что при любом $k > n_\varepsilon$ цикл z_k^r будет ε-гомологичен нулю в Φ.

Истинные циклы компакта Φ образуют группу $Z^r(\Phi)$, которая, между прочим, представляет собою, очевидно, топологический инвариант компакта Φ; в группе $Z^r(\Phi)$ имеем подгруппу $H^r(\Phi)$, состоящую из всех истинных циклов, гомологичных нулю. Фактор-группа $Z^r(\Phi)/H^r(\Phi)$ называется r-мерной группой Бетти $\Delta^r(\Phi)$ компакта Φ. Можно доказать, что это метрическое определение групп Бетти дает для компактов с точностью до изоморфизма тот же результат, что и определение при помощи нервов покрытий.

Третье обобщение групп Бетти есть обобщение чисто алгебраического характера. Исходной точкой построений, приведших нас к группам Бетти, было понятие цепи, т. е. линейной формы $\Sigma c_i t_i^r$ в ориентированных симплексах данного числа измерений с целочисленными коэффициентами. Однако можно рассматривать линейные формы не только с целочисленными коэффициентами, но и с коэффициентами, являющимися элементами любой наперед заданной группы A, которую мы будем называть *областью коэффициентов*. Получим «цепи по области коэффициентов A» и, отправляясь от них, будем строить все наши рассуждения совершенно так же, как в случае целочисленных цепей. В частности, для любой цепи $x^r = \Sigma c_i t_i^r$ по области коэффициентов мы снова определим границу x по формуле

$$\Delta x^r = \sum c_i \Delta t_i^r, \tag{3}$$

не смущаясь тем, что Δt_i^r есть целочисленная цепь

$$\Delta t_i^r = \sum \varepsilon_j^i t_j^{r-1} \qquad (\varepsilon_j = 0, +1, -1), \tag{4}$$

а коэффициенты c_i суть элементы группы: при подстановке (4) в (3) получим

$$\Delta x^r = \sum d_j t_j^{r-1}, \qquad d_j = \sum_i \varepsilon_j^i c_i,$$

что имеет совершенно ясный смысл, так как, по определению, $\varepsilon_j^i c_i$ есть c_i, $-c_i$, 0 в зависимости от того, будет ли $\varepsilon_j^i = 1$, -1 или 0. Все остальные наши рассуждения проходят дословно, как и раньше, и приводят нас к понятию r-мерной группы Бетти комплекса K по области коэффициентов A. Все это переносится и на случай компактов и на случай бесконечных комплексов (значит, и на случай открытых множеств эвклидовых пространств).

Среди групп, применяемых в качестве областей коэффициентов, кроме аддитивной группы всех целых чисел, наиболее важными являются следующие:

1° Конечные циклические группы; при этом циклическую группу порядка m удобнее всего представлять себе как группу целочисленных вычетов по модулю $m = 2, 3, \ldots$ Поэтому группы Бетти (а также цепи, циклы и т. п.), взятые по циклической группе порядка m, обычно называют группами Бетти и т. п. по модулю m.

2° Аддитивная группа R всех рациональных чисел.

3° Непрерывная циклическая группа $\varkappa$. Эту группу можно себе представлять, например, как аддитивную группу всех углов или всех вращений окружности вокруг ее центра. Но удобнее всего представлять себе непрерывную циклическую группу $\varkappa$ как фактор-группу аддитивной группы всех действительных чисел по подгруппе всех целых чисел (как «группа всех действительных чисел, приведенных по модулю 1»). Существенно заметить, что группа $\varkappa$ есть так называемая непрерывная или топологическая группа: множество всех ее элементов естественно рассматривать как топологическое пространство, гомеоморфное окружности.

4° Наконец, иногда бывает полезно рассматривать как область коэффициентов фактор-группу R_1 аддитивной группы всех рациональных чисел по подгруппе всех целых чисел (группа углов, соизмеримых с π, или вращений окружности на такие углы). Очевидно, эту группу можно рассматривать так же, как подгруппу группы $\varkappa$.

§ 6

ЗАКОН ДВОЙСТВЕННОСТИ АЛЕКСАНДЕРА — ПОНТРЯГИНА

В 1922 г. Александером была доказана следующая теорема, носящая его имя (теорема двойственности):

Если Φ *есть (кривой) полиэдр, лежащий в* R *и* $0 \leqslant r < n-1$, *то r-мерное число Бетти кривого полиэдра* Φ *равно* $(n-r-1)$*-мерному числу Бетти открытого множества* $\Gamma = R^n - \Phi$, *а нульмерное число Бетти открытого множества* Γ *(т. е. число областей, на которое* Γ *распадается) на единицу больше* $(n-1)$*-мерного числа Бетти* Φ.

Для $r = 0$ утверждение теоремы Александера почти очевидно: если нульмерное число Бетти полиэдра Φ равно s, то это значит, что этот полиэдр распадается на s связных полиэдров $\Phi_1, \Phi_2, \ldots, \Phi_s$. Беря разбиение пространства R^n на достаточно мелкие симплексы и отбирая те из них, которые имеют общие точки с данным Φ_i, можно (предполагая все время эти симплексы достаточно мелкими) заключить кривые полиэдры Φ_1, $\Phi_2, \ldots, \Phi_s$ в попарно непересекающиеся полиэдральные области, границы которых и дадут нам $(n-1)$-мерные циклы $z_1^{n-1}, \ldots, z_s^{n-1}$, лежащие в Γ; циклы эти будут *гомологически независимы между собою*, т. е. никакая нетривиальная линейная комбинация их не будет гомологична нулю в Γ. С другой стороны, всякий $(n-1)$-мерный цикл, лежащий в Γ, будет, как не трудно доказать, гомологичен некоторой линейной комбинации циклов $z_1^{n-1}, \ldots, z_s^{n-1}$. Отсюда и следует, что $(n-1)$-мерное число Бетти открытого множества Γ равно s.

В основном случае $1 \leqslant r \leqslant n-2$ теорема Александера сразу вытекает из следующей теоремы о зацепленных системах циклов, впервые доказанной в 1927 г. Л. С. Понтрягиным.

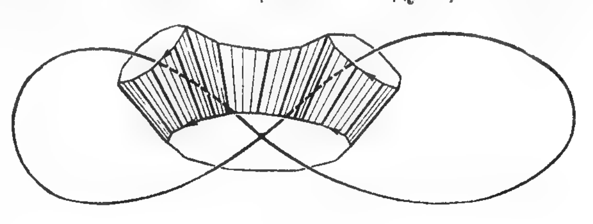
Пусть в Φ дана система гомологически независимых между собою циклов

$$z_1^r, \ldots, z_s^r$$

(это значит, что линейная комбинация $\sum c_i z_i^r$ этих циклов лишь тогда гомологична нулю в Φ, если все коэффициенты равны нулю). Тогда в Γ существует система циклов $z_1^{n-r-1}, \ldots, z_s^{n-r-1}$, «зацепленная» с си-

стемой в том смысле, что z_i^r и z_i^{n-r-1} зацеплены, но при $i \neq j$ циклы z_i^r и z_j^{n-r-1} не зацеплены. Обратно, если в Г дана система гомологически независимых между собою $(n-r-1)$-мерных циклов, то с этой системой зацеплена в только что указанном смысле некоторая система r-мерных циклов полиэдра Ф (фиг. 15).

Так как всякая система циклов в Ф (в Г), зацепленная с некоторой системой циклов в Г (в Ф), гомологически независима между собою в Ф (в Г), то из теоремы о зацепленных системах циклов следует, что число $(n-r-1)$-мерных гомологически независимых циклов в Г не меньше числа r-мерных гомологически независимых циклов в Ф и обратно, а это и значит, что r-мерное число Бетти Ф равно $(n-r-1)$-мерному числу Бетти Г.



Фиг. 15

Теорема о зацепленных системах циклов остается в силе и для $r=n-1$, но с тою оговоркою, что в обоих ее утверждениях рассматриваются среди нульмерных циклов лишь так называемые нормальные, т. е. такие, сумма коэффициентов которых равна нулю. Это условие вполне естественно: так как граница всякого одномерного симплекса $e_0\, e_1$ есть нормальный нульмерный цикл $e_1 - e_0$, то граница любой одномерной цепи также есть нормальный нульмерный цикл. Поэтому нульмерные циклы с суммой коэффициентов, отличной от нуля, не могут быть зацеплены ни с каким вообще циклом (так как на них, по только что замеченному, нельзя натянуть никакой пленки).

В силу сделанной нами оговорки, из теоремы о зацеплении в случае $r=n-1$ вытекает, что $(n-1)$-мерное число Бетти Ф равно максимальному числу гомологически независимых нормальных нульмерных циклов в Г. Но это последнее число, как легко видеть, на единицу меньше числа областей, на которые распадается открытое множество Г.

Доказательство теоремы о зацеплениях удобнее всего построить так. Для того случая, когда Ф есть полиэдр, сложенный из кубов некоторого разбиения пространства R^n, теорема доказывается комбинацией элементарно геометрических и чисто алгебраических соображений. Сведение

общего случая к этому элементарному можно осуществить, например, при помощи следующей леммы, доказанной мною в 1927 г. (уже после того, как Л. С. Понтрягин доказал свою теорему).

Лемма. Если Ф есть произвольный компакт, лежащий в R^n, то всякий истинный цикл компакта Ф, не гомологичный нулю в Ф, будет не гомологичным нулю и в некотором полиэдре, содержащем Ф и составленном из всех имеющих с Ф общие точки кубов некоторого достаточно мелкого разбиения пространства R^n. С другой стороны, каков бы ни был полиэдр Π_0, содержащий внутри себя множество Ф, можно найти лежащий в Π_0 и также содержащий внутри себя множество Ф полиэдр П, составленный из кубов некоторого разбиения пространства R^n, такой, что всякий истинный цикл этого полиэдра П гомологичен внутри Π_0 некоторому истинному циклу компакта Ф.

Однако из самого содержания этой леммы видно, что она позволяет свести к элементарному случаю не только случай кривого полиэдра, но и случай *любого компакта*, лежащего в R^n. Таким образом, мы убеждаемся в том, что формулировка теоремы о зацепленных системах циклов, а следовательно и формулировка закона двойственности Александера остаются в силе, если в них под Ф понимать любой лежащий в R^n компакт (в этом общем случае числа Бетти Ф, а значит, и Г могут, однако, оказаться бесконечными, что на утверждение теоремы Александера не оказывает никакого влияния).

Замечание. В теореме Александера в ее первоначальной формулировке для полиэдров, очевидно, содержится n-мерная теорема Жордана; в формулировке теоремы для любых компактов, естественно, содержится общее решение проблемы о разбиении пространства: число областей, на которые компакт Ф разбивает пространство R^n, на единицу больше мерного числа Бетти этого компакта и не меняется, если компакт Ф заменить любым гомеоморфным ему компактом Ф'.

После того как было установлено равенство *чисел* Бетти соответствующих размерностей Ф и Г, естественно возникает вопрос о связи между соответствующими группами Бетти. Если за область коэффициентов принять группу всех рациональных чисел, то изоморфизм групп Бетти сразу следует из равенства соответствующих чисел Бетти. Но как обстоит дело с другими областями коэффициентов и прежде всего с самой простой из них — с областью всех целых чисел?

Оказывается, что даже в случае, когда Ф — полиэдр, нельзя, вообще говоря, утверждать, что r-мерная группа Бетти Ф изоморфна $(n-r-1)$-мерной группе Бетти Г.

Именно, оказывается, что r-мерная группа кручения Ф изоморфна не $(n-r-1)$-мерной, а $(n-r-2)$-мерной группе кручения дополнительного множества Г. Картина, обнаружившаяся в случае полиэдров, совершенно запутывается, если перейти к любым компактам: группы Бетти компактов, вообще говоря, не имеют конечного числа образующих и поэтому не разлагаются в прямую сумму конечной группы кручения и определяемой числом Бетти «свободной группы». Это положение вещей не давало надежд даже на решение соответствующей проблемы инвариантности, а именно на доказательство следующей гипотезы: *Если* Ф и Ф' — *гомеоморфные компакты, лежащие в* R^n, *то целочисленные группы Бетти одного и того же числа измерений дополнительных множеств* $\Gamma = R^n - \Phi$ *и* $\Gamma' = R^n - \Phi'$ *изоморфны между собою*. Эта проблема и тесно связанная с нею проблема выражения групп Бетти открытого множества Г через какие-нибудь топологические инварианты дополнительного замкнутого множества Ф была блестящим образом решена в 1932 г. Л. С. Понтрягиным путем получения совершенно новой формулировки закона двойственности Александера, сохраняющей силу для любых областей коэффициентов. По этому поводу Л. С. Понтрягин построил одну из самых глубоких и значительных из математических теорий, созданных в последние десятилетия, именно так называемую теорию *характеров*.

Пусть дана группа G. Ее характером называется любое гомоморфное отображение этой группы в непрерывную циклическую группу $\varkappa$, причем если сама группа G непрерывная (т. е. является топологическим пространством, в котором сумма двух элементов непрерывно зависит от обоих этих элементов), то рассматриваются лишь непрерывные гомоморфизмы. Характеры данной группы G естественно образуют также группу, причем если группа G есть непрерывная компактная группа, то ее группа характеров определяется как дискретная группа; если же группа G дискретна, то ее группа характеров некоторым естественным образом топологизируется и в этой топологизации оказывается компактной. Эти определения приводят к следующей замечательной теореме: *Какова бы ни была компактная,*[1] *или счетная дискретная группа* G, *если* G^* — *группа*

[1] Со счетной базой.

характеров группы G, то группа характеров группы G^* есть снова группа G. Таким образом, каждая группа рассматриваемого нами класса (состоящего из компактных групп и из счетных дискретных) не только однозначно определяет свою группу характеров, но и сама однозначно ею определяется. Две группы называются *двойственными* между собою в смысле Понтрягина, если каждая из них есть группа характеров другой. На основании сказанного из двух двойственных между собою групп одна — всегда счетна и дискретна, а другая — компактна. Двойственными между собою являются, например, аддитивная группа всех целых чисел (бесконечная циклическая группа) и непрерывная циклическая группа $\varkappa$. Что касается конечных коммутативных групп, то каждая из них двойственна самой себе.

Теперь мы можем сформулировать и основное топологическое открытие Понтрягина:

Пусть Φ — компакт, лежащий в n-мерном эвклидовом пространстве R^n; пусть $\Gamma = R^n - \Phi$ — дополнительное к Φ открытое множество. Пусть A и A^* — две двойственные между собою группы, из которых A дискретна, а A^* компактна. Рассмотрим r-мерную группу Бетти $\Delta^r(\Phi, A^*)$ компакта Φ по области коэффициентов A^* и $(n - r - 1)$-мерную группу Бетти $\Delta^{n-r-1}(\Gamma, A)$ открытого множества Γ по области коэффициентов A. При надлежащей естественной топологизации группа $\Delta^r(\Phi, A^*)$ оказывается компактной и двойственной группе $\Delta^{n-r-1}(\Gamma, A)$.

В частности, если A — группа всех целых чисел, то A^* — непрерывная циклическая группа $\varkappa$, и мы видим, что $(n - r - 1)$-мерная группа Бетти открытого множества Γ по целочисленной области коэффициентов есть не что иное, как группа характеров r-мерной группы Бетти компакта Φ по группе $\varkappa$ как области коэффициентов.

Отсюда следует, что *целочисленные группы Бетти открытых множеств Γ, дополнительных к компактам Φ, инвариантно связаны с группами Бетти компактов Φ, но связь эта заключается не в изоморфизме, а в двойственности соответствующих групп.*

Пусть теперь Φ является, например, кривым полиэдром, а область коэффициентов есть конечная, например циклическая, группа. Тогда все группы Бетти, как Φ, так и $\Gamma = R^n - \Phi$, оказываются конечными, и все двойственности переходят в изоморфизмы.

Мы видим, к каким грандиозным обобщениям приходит, отправляясь

от такой простой вещи, как теорема Жордана, топология, и как в этих обобщениях простая геометрическая идея зацепления (остающаяся основой и понтрягинского закона двойственности) переплетается с широкими алгебраическими концепциями понтрягинской теории характеров, общематематическая сила и область приложимости которой выходят далеко за пределы топологии, но *поводом* к созданию которой послужила *задача обобщения* теоремы двойственности Александера (на случай замкнутых множеств вместо рассмотренных самим Александером кривых полиэдров). Это возникновение одного из значительнейших достижений современной математической мысли, каким является теория характеров, из стремления придать геометрическому факту его полную теоретико-множественную общность, поучительно с точки зрения закономерностей в развитии математических наук.

Я не могу здесь касаться связей, существующих между теорией характеров и другими областями современной математической мысли (в частности, связей с функциональным анализом). Скажу лишь несколько слов о дальнейшем развитии самой теории характеров и ее топологических приложений. Прежде всего, расширился самый класс групп, для которой строится теория характеров: класс счетных дискретных групп и компактных групп со счетной базой расширился сначала до любых дискретных и любых бикомпактных, а затем — до любых локально-бикомпактных групп; переход от данной локально-бикомпактной группы к ее группе характеров осуществляет взаимно однозначное отображение класса всех локально-бикомпактных групп на себя, при котором бикомпактные группы переходят в дискретные и обратно, причем группы со счетной базой переходят в группы со счетной базой. Такова наиболее общая форма понтрягинского алгебраического закона двойственности.

Что касается его топологических применений, то в работах А. Н. Колмогорова, Александера и моих они также чрезвычайно расширились и углубились. В частности, в моей работе «О гомологических свойствах расположения комплексов и замкнутых множеств»[1] удалось очень полным образом исследовать связи, существующие между группами Бетти замкнутого множества Φ и его дополнения $\Gamma = R - \Phi$ в очень широких предположениях относительно объемлющего пространства R. В частности,

[1] Известия АН СССР, серия математическая, том 6, вып. 5 (1943).

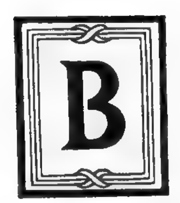
вполне исчерпан вопрос в том случае, когда R есть лкбой полиэдр. Что же касается случая, когда R — эвклидово пространство и вообще любое открытое многообразие, то нет надобности предполагать, что Ф — компакт: Ф может быть и неограниченным замкнутым множеством.

В самые последние месяцы мне удалось достигнуть нового существенного продвижения в этой области, а именно доказать *общий закон двойственности для любых (незамкнутых) множеств, лежащих в n-мерном пространстве.* Этот общий результат впервые включает в себя решение вопроса о гомологической двойственности элементарно геометрических фигур, не являющихся замкнутыми множествами. Я имею в виду множества, получающиеся из самых обыкновенных выпуклых многогранников применением в конечном числе операций сложения и *вычитания.* Для таких множеств (*а также для их топологических образов*) решение проблемы двойственности представляет непосредственный, наглядно геометрический интерес. Любопытно отметить, что это решение удалось получить, лишь подвергнув предварительно проблему двойственности самому широкому обобщению.

Отметим в заключение, что алгебраические методы в топологии позволили построением так называемой гомологической теории размерности совершенно по-новому осветить и понятие размерности, установив, в частности, глубокие связи, существующие между понятием размерности и законами двойственности типа Александера — Понтрягина.

Член-корреспондент АН СССР
Л. А. Люстерник

ТЕОРЕМА О ТРЕХ ГЕОДЕЗИЧЕСКИХ



1927—1929 гг. в работах Л. Г. Шнирельмана и автора настоящей статьи были развиты топологические методы качественного анализа вариационных задач. Наиболее интересным геометрическим приложением этих методов было решение известной задачи «о трех геодезических», т. е. доказательство существования трех замкнутых геодезических на выпуклых поверхностях.[1] Была доказана более общая теорема — доказательство существования трех замкнутых и самонепересекающихся геодезических на поверхностях рода ноль, т. е. поверхностях, которые можно взаимно однозначно и непрерывно отобразить на сферу. Задачи о замкнутых геодезических были поставлены Пуанкаре в связи с исследованием периодических движений. Замкнутые геодезические на поверхностях отвечают периодическим движениям с двумя степенями свободы.

Полная формулировка доказанной теоремы такова.

На всякой замкнутой трижды дифференцируемой поверхности S_2 *рода ноль существуют:*

---

[1] Л. Люстерник и Л. Шнирельман. Топологические методы в вариационных задачах. Тр. Ин-та матем. и мех. МГУ, 1930 (отд. выпуск). Новую редакцию этого мемуара см. Усп. матем. наук, 1947, т. 2, вып. 1 (17).

а) *или три замкнутые самонепересекающиеся геодезические разной длины,*

б) *или покрывающее S_2 семейство таких геодезических равной длины и замкнутая самонепересекающаяся геодезическая отличной длины,*

в) *или все геодезические на S_2 являются замкнутыми самонепересекающимися кривыми равной длины.*

Если S_2 эллипсоид, то эти геодезические суть главные сечения эллипсоида; случаи а), б), в) бывают тогда, когда есть трехосный эллипсоид, эллипсоид вращения или сфера соответственно. Можно вместо S_2 рассматривать сферу S_0, на которой за расстояния между точками мы принимаем расстояния между образами этих точек на S_2.

Рассмотрим на S_0 три семейства линий:

1) семейство A_3 всех окружностей,

2) семейство A_2 всех больших кругов,

3) семейство A_1 больших кругов с общим диаметром.

Рассмотрим на S_0 непрерывные деформации кривых и семейств кривых (причем непрерывная деформация семейства сводится к непрерывной деформации всех его кривых, при которой достаточно близкие кривые остаются сколь угодно близкими по положению и длине). Обозначим через $[A_i]$ $(i=1, 2, 3)$ класс всех семейств спрямляемых кривых, получаемых на S_0 непрерывной деформацией семейства A_i. Обозначим через $c(A)$ максимум длин кривых семейства A, через c_i—нижнюю границу чисел $c(A)$ для всех A из $[A_i]$.

Имеем

$$c_3 \geqslant c_2 \geqslant c_1 > 0.$$

Последнее неравенство вытекает из того, что однопараметрическое семейство A_1 покрывает сферу и это свойство сохраняется у всех семейств A из $[A_1]$, а поэтому соответственные числа $c(A)$ не могут быть меньше некоторой положительной константы.

Теорема 1. *Если A_i^0 есть семейство класса $[A_i]$, для которого $c(A_i^0)=c_i$ $(i=1, 2, 3)$, то A_i^0 содержит замкнутую геодезическую длины c_i.*

Из теоремы 1 при условии $c_3 > c_2 > c_1$ следует случай а) основной теоремы.

Теорема 1 доказывается при помощи операции «стягивания», т. е. деформации дуги $\breve{ab}$ в минимальную геодезическую дугу $\overline{ab}$, соединяющую те же точки a и b. Если q не есть замкнутая геодезическая, то стягиванием ее дуги можно уменьшить длину q. Операцию стягивания, примененную к кривой, можно распространить на близкие кривые. Если в семействе A_i^0 все кривые длины c_i отличны от замкнутых геодезических, то можно определить деформацию D семейства A_i^0, сводящуюся к стягиванию дуг его кривых, при которой длины кривых из A_i^0 уменьшаются и семейство A_i^0 деформируется в семейство A_i^1 того же класса A_i, причем $c(A_i^1) < c(A_i^0) = c_1$; это противоречит определению c_1, и теорема доказана.

Можно не предполагать существования семейства A_i^0. Во всяком случае существует семейство A_i^ε класса $[A_i^0]$, на котором $c(A_i^\varepsilon) < c_i + \varepsilon$ при любом $\varepsilon > 0$. Тогда аналогичными рассуждениями можно при достаточно малом ε доказать существование «почти-геодезической» (т. е. кривой, дуги которой весьма мало отличаются от геодезических), длина которой сколь угодно близка к c_i. Предельным переходом при $\varepsilon \to 0$ доказывается существование замкнутой геодезической длины c_i.

Пусть теперь $c_i = c_2 = c$. Допустим для простоты, что существует семейство A_2^0 класса $[A_2]$, для которого $c(A_2^0) = c_2 = c$.

В силу теоремы 1 в A_2^0 существует непустое множество P замкнутых геодезических длины $c_2 = c$. Имеет место следующее утверждение: $A_2^0 - P$ *не содержит ни одного множества класса* $[A_1]$. Допустим, наоборот, что в $A_2^0 - P$ входит множество A_1^0 класса $[A_1]$. Тогда, с одной стороны, $c(A_1^0) \leqslant c(A_2^0) = c = c_1$; с другой стороны, $c(A_1^0) \geqslant c_1$ по определению c_1. Поэтому $c(A_1^0) = c_1$; но тогда в силу теоремы 1 существует в A_1^0 замкнутая геодезическая q длины $c_1 = c$; q должна входить в P, между тем q входит вместе со всем A_1^0 в $A_2^0 - P$. Это противоречие доказывает наше утверждение.

Итак, при $c_1 = c_2 = c$ семейство P замкнутых геодезических длины c должно быть настолько богато, чтобы $A_2^0 - P$ не содержало множества из $[A_1]$.

Вернемся к исходным множествам A_3, A_2, A_1. Если каждую окружность принять за точку и идентифицировать все вырожденные окружности радиуса O, то A_3 превратится в проективное трехмерное пространство, A_2 — в проективную плоскость, A_1 — в проективную прямую. В абстрактном пространстве L всех спрямляемых замкнутых кривых на S_0 семейства A_i суть проективные пространства i измерений.

Если A_2^0 есть проективная плоскость в L, то сразу видно, что множество P, для которого $A_2^0 - P$ не содержит (топологической) проективной прямой из $[A_1]$, должно быть достаточно обширно («высокой категории»); оно должно содержать некий континуум[1] и даже больше: грубо говоря, оно само должно содержать проективную прямую.

Вообще же A_2^0 получается деформацией проективной плоскости A_2 в L, но при этом оно сохраняет свойства разбиения A_2^0 на части: если часть $A_2^0 - P$ (дополнительная к P) не содержит «проективной прямой» из $[A_1]$, то P «достаточно богато». Совокупность P (замкнутых геодезических длины c) должна содержать континуальную систему элементов; более детальный анализ показывает, что совокупность замкнутых геодезических длины c покрывает S_0 (случай б) основной теоремы).

То же самое мы получим при $c_1 < c_2 = c_3$.

При условии $c_1 = c_2 = c_3$ из свойств разбиения проективного трехмерного пространства на части (так называемые категории этого пространства) следует случай в) основной теоремы.

Теорема о существовании трех геодезических в ослабленной формулировке (без предположения самонепересекаемости) основана, таким образом, на некоторых топологических свойствах разбиения проективных пространств (так называемых «категорий» проективных пространств[2]) и их расположения в абстрактном пространстве L и на некоторых рассмотрениях, общих для разных вариационных задач. Но при этом одна из геодезических может выродиться в k раз повторенную другую,

[1] Это вытекает из свойства проективной плоскости: она не может быть разбита на две части, сводимые к точке («категория» проективной плоскости больше двух).

[2] Проективное пространство n измерений имеет «категорию» $n+1$, т. е. может быть разбито на $n+1$, но не может быть разбито на n частей, сводимых к точке.



Л. Шнирельман

184 v 185

причем $k > 1$. Чтобы исключить такой случай, мы сохраняем в классах $[A_i]$ только семейства, состоящие из самонепересекающихся кривых.

Если c_i обозначают нижние грани $c[A]$, где A — семейства самонепересекающихся кривых из $[A_i]$, то теорема 1 уточняется: она будет говорить о существовании замкнутой самонепересекающейся геодезической длины c_i.

Теперь уже нельзя пользоваться операцией стягивания дуги $\breve{ab}$ самонепересекающейся кривой q к минимальной геодезической $\overline{ab}$, так как при этом кривая q может стать самопересекающейся. Но можно, оказывается, провести дополнительное преобразование полученной кривой q_1 так, что кривая q_1 перейдет опять в самонепересекающуюся кривую меньшую по длине, чем q. Всюду в доказательстве теоремы 1 следует заменить операцию стягивания этой более сложной операцией; тогда доказательство существования замкнутой геодезической перейдет в доказательство существования самонепересекающейся замкнутой геодезической, и основная теорема докажется целиком.

Для выпуклых поверхностей, как показал Погорелов, можно теорему о существовании трех замкнутых геодезических обобщить на нерегулярные поверхности.

Для n-мерного случая автор настоящей статьи сравнительно недавно доказал, при некоторых метрических ограничениях, существование на всяком многообразии, гомеоморфном n-мерной сфере, $n+1$ замкнутых геодезических и $2n-1$ замкнутых геодезических при $n=2^k$, причем или длины этих геодезических различны, или появляется континуальное семейство замкнутых геодезических.[1]

---

[1] Доказательство приводится в мемуаре автора. Тр. Матем. ин-та имени В. А. Стеклова АН СССР, т. XIX.

Академик
В. И. Смирнов

РАБОТЫ В. А. СТЕКЛОВА О РАЗЛОЖЕНИЯХ ПО ОРТОГОНАЛЬНЫМ ФУНКЦИЯМ

1. В своей научной работе Владимир Андреевич Стеклов уделял больше всего внимания различным вопросам математической физики, и в этом надо искать источник его многочисленных исследований о разложениях по ортогональным функциям. После блестящих работ в первой половине XIX века в области математической физики наступило сравнительное затишье. Критические исследования обнаружили ряд недочетов в отношении математической строгости в упомянутых выше работах. Последнее десятилетие XIX века было поворотным пунктом в истории математической физики: с этого времени начали создаваться новые методы, с помощью которых не только были устранены указанные недочеты, но и открыты совершенно новые перспективы для решения основных задач математической физики. Одним из первых в этом новом периоде развития математической физики был В. А. Стеклов. Укажем историю тех вопросов, которые привели В. А. Стеклова к работам о разложениях по ортогональным функциям; ограничимся при этом лишь простейшими задачами.

Задача о распространении тепла в неоднородном стержне может быть приведена к уравнению с частными производными

$$p(x)\frac{\partial v}{\partial t}=\frac{\partial^2 v}{\partial x^2}-q(x)\,v, \qquad (1)$$

где $v(t, x)$—температура стержня в его точке x и в момент времени t, а $p(x)$ и $q(x)$ — известные функции x. Мы считаем стержень ограниченным, так что $a \leqslant x \leqslant b$, функции $p(x)$ и $q(x)$ — непрерывными и положительными в этом промежутке. Кроме уравнения (1) должно быть задано начальное распределение температуры

$$v_{t=0} = f(x) \tag{2}$$

и какие-либо предельные условия на концах, например:

$$\frac{\partial v}{\partial x} - hv = 0 \quad \text{при} \quad x = a; \quad \frac{\partial v}{\partial x} + Hv = 0 \ \text{при} \ x = b, \tag{3}$$

где h и H — положительные постоянные, или

$$v|_{x=a} = 0, \qquad v|_{x=b} = 0. \tag{3'}$$

Фурье и Пуассон создали общий метод решения задач указанного типа, состоящий в следующем.

Ищем решение уравнения (1) в виде произведения $v = e^{-\lambda t} u(x)$. Для $u(x)$ получаем обыкновенное дифференциальное уравнение

$$u''(x) + [\lambda p(x) - q(x)]\, u = 0, \tag{4}$$

и, кроме того, в силу (3) или (3′), искомая функция $u(x)$ должна удовлетворять предельным условиям

$$u'(a) - h\,u(a) = 0, \qquad u'(b) + Hu(b) = 0, \tag{5}$$

или

$$u(a) = u(b) = 0. \tag{5'}$$

Уравнение (4) при условиях (5) или (5′) имеет очевидное решение $u(x) = 0$. Но, как оказывается, существует бесчисленное множество $\lambda_1, \lambda_2, \ldots$ вещественных положительных значений параметра λ, при которых задача (4), (5) или (4), (5′) имеет решения, отличные от нуля. При всех других значениях λ, отличных от указанных выше, задача имеет только нулевое решение. Обозначим через $u_k(x)$ ненулевое решение задачи при $\lambda = \lambda_k$. Ввиду линейности и однородности (4) и (5) можно умножить $u_k(x)$ на произвольную постоянную A_k. Таким образом, уравнение (1) имеет бесчисленное множество решений $A_k e^{-\lambda_k t} u_k(x)$,

удовлетворяющих предельным условиям (3) или (3'). Ввиду линейности и однородности (1) и (3) и сумма

$$v(t, x)=\sum_{k=1}^{\infty} A_k e^{-\lambda_k t} u_k(x)$$

будет удовлетворять уравнению (1) и предельным условиям. Остается подобрать постоянные A_k так, чтобы удовлетворялось начальное условие (2). Это приводит к равенству

$$f(x)=\sum_{k=1}^{\infty} A_k u_k(x), \qquad (6)$$

т. е. к разложению заданной функции $f(x)$ по функциям $u_k(x)$.

Совершенно аналогична задача о распространении тепла в трехмерном однородном теле D, ограниченном поверхностью S. Она сводится к решению дифференциального уравнения

$$\frac{\partial^2 v}{\partial t^2}=a^2 \Delta v, \qquad (7)$$

где a^2 — постоянная, причем

$$\Delta v=\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2};$$

при начальном условии

$$v|_{t=0}=f(x, y, z)$$

и предельном условии

$$\frac{\partial v}{\partial n}+hv=0 \quad \text{на } S \qquad (8)$$

или

$$v=0 \text{ на } S, \qquad (8')$$

где n—направление внешней нормали к S. В дальнейшем для сокращения письма мы будем зачастую писать $f(M)$, вместо $f(x, y, z)$, понимая под (x, y, z) координаты точки M. Решение уравнения (7) ищется в виде: $v(t; M)=e^{-Mt}u(M)$. Для $u(M)=u(x, y, z)$ получаем уравнение

$$\Delta u+\lambda u=0 \quad \left(\lambda=\frac{\mu}{a^2}\right), \qquad (9)$$

и предельное условие

$$\frac{\partial u}{\partial n}+hu=0 \quad \text{на } S \tag{10}$$

или

$$u=0 \quad \text{на } S. \tag{10'}$$

Существует бесчисленное множество $\lambda_1, \lambda_2, \ldots$ положительных значений λ, при которых эта задача имеет решения $u_1(M), u_2(M), \ldots$, отличные от нулевого. Решение основной задачи ищется в виде

$$v(t; M)=\sum_{k=1}^{\infty} A_k e^{-a^2\lambda_k t} u_k(M) \tag{11}$$

и начальное условие приводит к задаче разложения заданной функции $f(M)$ по функциям $u_k(M)$

$$f(M)=\sum_{k=1}^{\infty} A_k u_k(M). \tag{12}$$

Таким образом, метод Фурье — Пуассона приводит к следующим двум основным задачам:

1) Доказательство существования и исследование свойств «характеристических чисел» λ_k и соответствующих им «фундаментальных функций» u_k.

2) Исследование возможности разложения произвольной заданной функции f по фундаментальным функциям.

В соответствии с целью, поставленной в настоящей статье, мы будем говорить почти исключительно о второй задаче.

2. Наиболее подробное исследование этой задачи в первой половине XIX века принадлежит Лиувиллю. Он исследовал эту задачу (1837 г.) в применении к уравнению (4). Однако и его исследования лишены надлежащей строгости. До конца XIX века, в течение шестидесяти лет, не было никаких успехов в решении поставленной задачи. В 1896 г. В. А. Стеклов исследовал задачу о разложении при некоторых ограничительных условиях довольно общего характера для уравнений (4) и (9), и с этого времени и до конца жизни он печатал большое число замечательных работ по этому вопросу. В краткой статье мы сможем дать лишь общий очерк его идей и результатов.

Не трудно показать, что фундаментальные функции уравнения (4) обладают следующим свойством «ортогональности с весом $p(x)$»:

$$\int_a^b p(x)\,u_k(x)\,u_l(x)\,dx = 0 \quad \text{при } k \neq l; \tag{13}$$

и, кроме того, умножая $u_k(x)$ на постоянный множитель, можно считать эти функции «нормированными», что выражается равенством

$$\int_a^b p(x)\,u_k^2(x)\,dx = 1 \qquad (k = 1, 2, \ldots). \tag{14}$$

При этом коэффициенты в предлагаемом разложении (6) естественно определить по заданной функции $f(x)$ равенствами

$$A_k = \int_a^b p(x)\,f(x)\,u_k(x)\,dx. \tag{15}$$

Совершенно аналогично для фундаментальных функций уравнения (9) мы получим

$$\iiint\limits_D u_k(M)\,u_l(M)\,d\tau = \begin{cases} 0 & \text{при } k \neq l, \\ 1 & \text{при } k = l \end{cases},$$

и в предполагаемом разложении (12)

$$A_k = \iiint\limits_D f(M)\,u_k(M)\,d\tau, \tag{16}$$

где $d\tau = dx\,dy\,dz$ — элемент объема.

Начнем с выяснения одной общей, чрезвычайно ценной идеи, впервые выдвинутой Стекловым и широко использованной им при решении задачи разложений. Пусть $u_k(x)$ — какая угодно ортогональная, нормированная система с весом $p(x)$, т. е. удовлетворяющая равенствам (13) и (14). Мы считаем $u_k(x)$ непрерывными, функцию $p(x)$ неотрицательной. Функция $p(x)$ может обращаться в отдельных точках в нуль и терять непрерывность, но так, что она остается интегрируемой. Пусть $f(x)$ — любая интегрируемая функция, которую мы будем пока считать ограниченной. Составим по формулам (15) ее «коэффициенты Фурье A_k

относительно ортогональной нормированной системы $u_k(x)$». Вводя обозначение погрешности $R_n(x)$ в приближенном представлении функции $f(x)$ «отрезком ее ряда Фурье по функциям $u_k(x)$»

$$R_n(x)=f(x)-\sum_{k=1}^{n} A_k u_k(x), \tag{17}$$

получаем следующее простое следствие формул (13), (14) и (15):

$$\int_a^b p(x)\, R_n^2(x)\, dx=\int_a^b p(x)\,[f(x)]^2\, dx-\sum_{k=1}^{n} A_k^2. \tag{18}$$

Отсюда непосредственно видно, что при возрастании n сумма, стоящая справа, не может оказаться больше интеграла от $p(x)[f(x)]^2$, и, таким образом, обязательно сходится ряд с неотрицательными членами

$$\sum_{k=1}^{\infty} A_k^2. \tag{19}$$

Интеграл, стоящий в левой части уравнения (18), дает среднюю квадратичную погрешность в приближенном представлении $f(x)$ отрезком ее ряда Фурье. Он не может увеличиваться при возрастании n. Стремление этой квадратичной погрешности к нулю при беспредельном возрастании n равносильно, очевидно, тому, что сумма ряда (19) равна в точности $\int_a^b p(x)\,[f(x)]^2\, dx$, т. е. уменьшаемому в правой части (18). Итак, равенство

$$\int_a^b p(x)\,[f(x)]^2\, dx=\sum_{k=1}^{\infty} A_k^2 \tag{20}$$

равносильно равенству

$$\lim_{n\to\infty}\int_a^b p(x)\, R_n^2(x)\, dx=0. \tag{21}$$

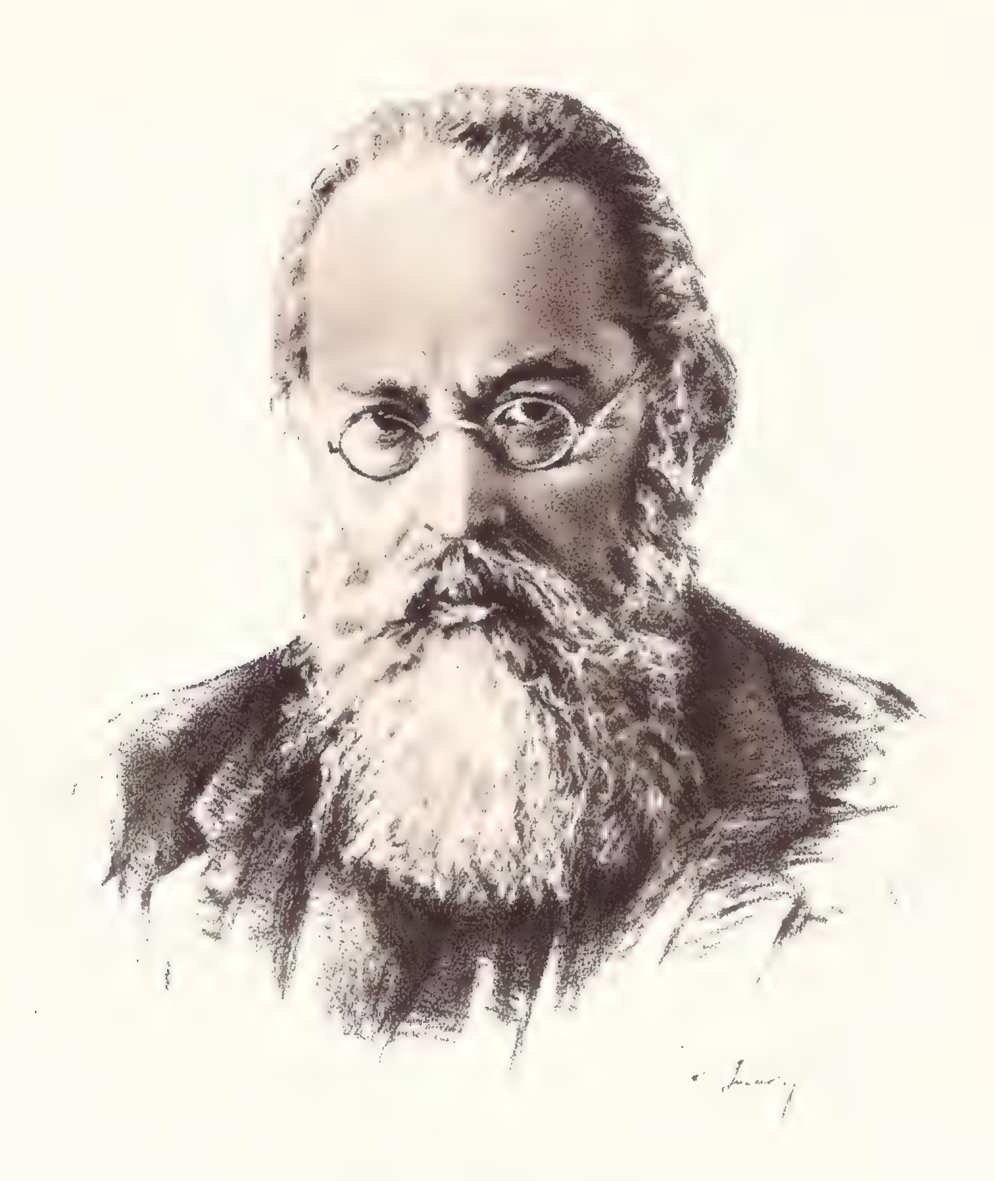
Равенство (20) представляет собою теорему Пифагора для пространства функций $f(x)$, в котором $u_k(x)$ приняты за орты. Соблюдение (20) для

любой функции $f(x)$ равносильно по существу тому, что систему ортов $u_k(x)$ мы не можем пополнить новыми ортами. В трехмерном пространстве наличие трех ортов уже гарантирует, очевидно, их полноту. В бесконечномерном пространстве функций мы не имеем возможности путем простого счета ортов удостовериться в их полноте, и доказательство (20) для любой $f(x)$ составляет первую и основную проблему теории ортогональных систем функций. Стеклов первый поставил в общем виде эту проблему и исследовал ее, о чем мы будем говорить ниже. Отметим, что, вместо ограниченности $f(x)$, достаточно требовать существование интеграла $\int\limits_a^b p(x)[f(x)]^2 dx$. Все сказанное выше годится и для случая функций от многих переменных, в частности и для фундаментальных функций уравнения (9).

3. Доказательству равенства (20) при любом выборе $f(x)$ для различных систем функций u_k посвящено большое число работ Стеклова, начиная с первых работ 1896 г. С 1910 г. Стеклов называет равенство (20) *условием замкнутости*, а системы функций u_k, для которых оно соблюдено при любом выборе $f(x)$, *замкнутыми системами*.

При доказательстве формулы замкнутости Стеклов обнаруживает следующий общий факт. Если эта формула доказана для некоторого специального класса функций, то отсюда необходимо следует, что она справедлива для всех функций $f(x)$, для которых интеграл $\int\limits_a^b p(x)[f(x)]^2 dx$ существует. В случае конечного промежутка (a, b), о чем мы и говорили раньше, или при некоторых дополнительных условиях относительно $p(x)$ таким специальным классом может быть класс полиномов. Таким образом, специфичность системы функций u_k используется лишь при доказательстве формулы замкнутости для упомянутого выше специального класса. Если $u_k(x)$ есть полином степени k, то замкнутость относительно полиномов очевидна.

Далее Стеклов изучил свойства замкнутых систем. Приведем два основных результата, полученных им в этом направлении. Они имеют существенное значение при исследовании теорем разложения. Первый состоит в следующем.



В. Стекловъ

Если ряд Фурье непрерывной функции

$$\sum_{k=1}^{\infty} A_k u_k(x) \tag{22}$$

по замкнутой ортогональной системе $u_k(x)$ равномерно сходится в основном промежутке (a, b) или какой-либо его части (a_1, b_1), то его сумма равна $f(x)$ в этом промежутке равномерной сходимости.

Второй результат относится к общему случаю, когда ряд Фурье (22) может и не быть вовсе сходящимся. Стеклов доказывает следующее.

Если система $u_k(x)$ замкнута и (22) есть ряд Фурье $f(x)$, то для любой функции $\varphi(x)$ такой, что интеграл $\int_{a_1}^{b_1} p(x)[\varphi(x)]^2\,dx$ существует, справедливо уравнение

$$\int_{a_1}^{b_1} p(x) f(x) \varphi(x)\,dx = \sum_{k=1}^{\infty} A_k \int_{a_1}^{b_1} p(x) \varphi(x) u_k(x)\,dx, \tag{23}$$

где (a_1, b_1) — любая часть промежутка (a, b). Таким образом, оказывается, что даже если ряд Фурье (22) расходится, то его почленное интегрирование приводит к сходящемуся ряду, и сумма этого ряда дает интеграл от левой части. Если, например, $p(x) \equiv 1$, и мы положим $\varphi(x) \equiv 1$, то

$$\int_{a_1}^{b_1} f(x)\,dx = \sum_{k=1}^{\infty} A_k \int_{a_1}^{b_1} u_k(x)\,dx. \tag{24}$$

Этот результат, кроме чисто практических его следствий при изучении свойств разложений, имеет и принципиальное значение. Он доказывает, что хотя бесконечный ряд (22) может и не давать разлагаемой функции $f(x)$ в отдельных точках x, но, проинтегрированный почленно, он в точности дает средние от функции по промежутку (интеграл). В дальнейшем эта идея осреднения, приведшая к понятиям функции промежутка, области или множества, вошла в аппарат современной математической физики и сблизила его с реальными условиями физического эксперимента. Прежде понятие функции точки было явно абстрактным по отношению

к эксперименту. Указанные выше два основных результата теории замкнутости почти непосредственно следуют из определения замкнутых систем.

Отметим, в частности, что если (a_1, b_1) совпадает со всем промежутком (a, b), то (23) принимает вид

$$\int_a^b p(x) f(x) \varphi(x) dx = \sum_{k=1}^{\infty} A_k B_k, \tag{25}$$

где A_k — коэффициенты Фурье $f(x)$ и B_k — коэффициенты Фурье $\varphi(x)$. Уравнение (25) непосредственно следует из (20), а из (25) может быть получено уравнение (23), если взять $\varphi(x)$ равным нулю вне (a_1, b_1).

В указанных выше случаях параметр λ, определявший собственные значения предельной задачи, входил в дифференциальное уравнение. Стеклов рассматривал фундаментальные функции и в тех случаях, когда λ входит в предельное условие. Приведем определение так называемых «фундаментальных функций Стеклова».

Внутри трехмерной области D, ограниченной поверхностью S, должно быть выполнено уравнение

$$\Delta u = 0 \tag{26}$$

и на поверхности S предельное условие

$$\frac{\partial u}{\partial n} = \lambda p u,$$

где n — направление нормали к S и $p(M)$ — заданная на S положительная функция. Существует бесчисленное множество значений $\lambda_1, \lambda_2, \ldots$ параметра λ, при которых эта задача имеет решения $u_1(M), u_2(M), \ldots$, отличные от нулевого. Эти решения и называются *фундаментальными функциями Стеклова*. Они обладают ортогональностью на S с весом $p(M)$ и являются распространением на общие поверхности шаровых функций в случае сферы, функций Ляме — в случае эллипсоида. Известно основное значение такого рода функций при решении задачи Дирихле и Неймана для уравнения (26).

Другие способы построения аналогичных функций почти одновременно со Стекловым были указаны Пуанкаре и Ле-Руа. В мемуаре

1904 г. Стеклов указывает все системы функций, для которых им доказана замкнутость. Сюда вошли, между прочим, фундаментальные функции Стеклова и функции Ле-Руа. Отметим, что теория замкнутости применялась Стекловым не только к теоремам разложения, но и к ряду других вопросов математического анализа.

Не касаясь методов, которыми пользовался Стеклов в своих исследованиях по теории замкнутости, укажем, в связи с предыдущим, на один из них. Пусть $f(x)$ задана на промежутке (a, b) и продолжена во вне этого промежутка, например так, что $f(x) = f(a)$ при $x < a$ и $f(x) = f(b)$ при $x > b$. Составляем новую функцию

$$F(x) = \frac{1}{h} \int_{x}^{x+h} f(\xi)\, d\xi\,, \tag{25'}$$

где h — малое положительное число. Если $f(x)$ только интегрируема, то $F(x)$ непрерывна; если $f(x)$ только непрерывна, то $F(x)$ имеет непрерывную производную; если $f(x)$ имеет непрерывные производные до порядка k, то $F(x)$ имеет непрерывные производные до порядка $(k+1)$. С другой стороны, при непрерывности $f(x)$ разность $F(x) - f(x)$ становится сколь угодно малой по абсолютной величине, когда h стремится к нулю, т. е. сглаживание функции $f(x)$, согласно (25′), приводит к близкой функции $f(x)$ с лучшими свойствами. Операцию (25′) можно повторить несколько раз. Она применима и к функциям многих переменных. Начиная с 1910 г., Стеклов, используя этот метод сглаживания, показывает, что если условие замкнутости (20) имеет место для некоторого узкого класса функций, то оно имеет место и для более широкого класса. При наличии одной интегрируемости $f(x)$ Стеклов рассматривает те значения x, при которых правая часть уравнения (25′) имеет предел при стремлении h к нулю.

В своих общих исследованиях по теории замкнутости Стеклов пользовался классическим понятием интеграла, тем, что мы теперь называем интегралом Римана. В дальнейшем выяснилось, что эта теория Стеклова принимает законченный вид, если использовать новое понятие интеграла Лебега или Лебега — Стильтьеса.

4. Переходим к результатам работы Стеклова, касающимся разложения заданной функции по ортогональной системе $u_k(x)$. Вопрос здесь

13*

сводится к выяснению тех условий, при которых ряд Фурье (22) функции $f(x)$ сходится и его сумма равна $f(x)$:

$$f(x) = \sum_{k=1}^{\infty} u_k(x). \tag{26'}$$

Иначе говоря, остаточный член $R_n(x)$, определенный равенством (17), должен стремиться к нулю, что может иметь место или для отдельных значений x, или для всех x из промежутка (a, b). Стеклов занимается главным образом последним случаем. Напомним, что условие замкнутости равносильно стремлению $R_n^2(x)$ к нулю в среднем на промежутке (a, b).

В большинстве работ Стеклов сначала устанавливает замкнутость системы, а затем переходит к исследованию разложений и пользуется указанными выше свойствами замкнутых систем. Условия, при которых имеет место разложение (26′), касаются системы $u_k(x)$ и свойств разлагаемой функции $f(x)$. В случае фундаментальных функций уравнения (4) условия относятся к коэффициентам $p(x)$ и $q(x)$ уравнения и форме предельных условий.

Мы изложили сущность теоремы разложения для функций одного переменного. Все это относится и к функциям многих переменных, в частности к фундаментальным функциям уравнения (9). Первые работы Стеклова по теоремам разложения были опубликованы в 1896 г., а последние — в 1926 г. За эти тридцать лет Стеклов часто занимался вопросами разложений по одним и тем же ортогональным функциям; его главные усилия были направлены на то, чтобы получить наиболее широкие для $f(x)$ условия теоремы разложения. Кроме результатов теории замкнутости, он применял и другие методы, о чем будет сказано ниже. Приведем кратко некоторые результаты его работ.

В работе 1896 г. для разложения по фундаментальным функциям уравнения (4) при предельных условиях (5) требуется следующее: $p(x)$ и $q(x)$ — непрерывные положительные функции, $f(x)$ имеет производные до четвертого порядка, удовлетворяет предельным условиям (5), и, кроме того, функция $(fq - f') : p$ также удовлетворяет предельным условиям. В работе 1913 г. $p(x)$ считается непрерывной и положительной, $q(x)$ — только непрерывной, предельные условия берутся, например, в следующей общей форме:

$$u'(a) = \alpha u(a) + \beta u(b), \quad u'(b) = \gamma u(a) + \delta u(b),$$

где $\alpha + \delta = 0$. При этом доказывается уравнение (26′) и равномерная сходимость ряда на промежутке (a, b) для всех функций, удовлетворяющих условию Липшица

$$|f(x)-f(y)| \leqslant m|x-y|, \tag{27}$$

где m — некоторая постоянная. При предельных условиях (3′) придется добавить еще условия $f(a) = 0$ и $f(b) = 0$. Из приведенного примера видно, насколько удавалось улучшить условия применимости теоремы разложения. В некоторых работах Стеклов исследовал фундаментальные функции уравнения (9) при предельном условии (10′). В работе 1910 г. доказываются следующие наиболее широкие условия теоремы разложения: $f(M)$ имеет внутри D непрерывные производные до второго порядка и обращается в нуль на поверхности S.

Большое внимание Стеклов уделил разложению по полиномам, которые были введены еще Чебышевым (1855 г.) в связи с теорией непрерывных дробей. Если (a, b) — конечный промежуток и $p(x)$ — неотрицательная интегрируемая функция, то полиномы Чебышева с весом $p(x)$ удовлетворяют обычным условиям ортогональности и нормальности

$$\int_a^b p(x)\varphi_k(x)\varphi_l(x)dx = \begin{cases} 0 \text{ при } k \neq l, \\ 1 \text{ при } k = l. \end{cases} \tag{28}$$

При этом $\varphi_k(x)$ есть полином степени k. Условия (28) при заданном весе вполне определяют полиномы $\varphi_k(x)$ в предположении, что у каждого $\varphi_k(x)$ старший коэффициент есть положительное число. Разложениями по полиномам Чебышева Стеклов занимался с 1902 г. В большом цикле работ, опубликованном им уже в 1921 г., он дает свои наиболее совершенные результаты по этому вопросу. Пользуясь теорией замкнутости, он получает следующий результат.

Если $f(x)$ удовлетворяет на (a, b) условию Липшица (27) и вес имеет вид

$$p(x) = (x-a)^\alpha (b-x)^\beta q(x), \tag{29}$$

где α и $\beta > -1$, $q(x)$ — положительная и удовлетворяет условию (27), то $f(x)$ разлагается в равномерно сходящийся ряд Фурье по полиномам Чебышева во всяком промежутке (a_1, b_1), лежащем внутри (a, b). В том же 1921 г. Стеклов опубликовал два мемуара о разложениях по поли-

номам Чебышева, не зависимые от теории замкнутости. Эти мемуары были им приурочены к столетию со дня рождения Чебышева. Результаты этих работ идут больше в направлении облегчения условий, налагаемых на разлагаемую функцию $f(x)$. При общих предположениях о весе $p(x)$ Стеклов получает следующую оценку для полиномов Чебышева:

$$|\varphi_k(x)| \leqslant M\sqrt{k}, \tag{30}$$

где M — некоторая постоянная. При специальном выборе веса полиномы могут оставаться ограниченными при возрастании k, т. е.

$$|\varphi_k(x)| \leqslant M \quad (k=0, 1, 2, \ldots). \tag{31}$$

Стеклов доказывает, что если это условие выполняется в некоторой точке x, лежащей внутри (a, b), то всякая функция, удовлетворяющая общему условию Липшица

$$|f(x)-f(y)| \leqslant m\,|x-y|^{\gamma} \quad (0<\gamma\leqslant 1), \tag{32}$$

разлагается в упомянутой точке в ряд Фурье по полиномам Чебышева. Условие (31) выражено не через вес, и Стеклов замечает, что ему не удалось получить общих и простых условий для веса, которые гарантировали бы выполнение условия (31). Они выполняются во всякой точке внутри (a, b), например для полиномов Якоби, которым соответствует вес $p(x) = (x-a)^{\alpha}(b-x)^{\beta}$ при α и $\beta > -1$. Мы знаем теперь, что если вес $p(x)$ положителен и удовлетворяет условию Липшица (27), то $|\varphi_k(x)| \leqslant M[(x-a)(b-x)]^{-\frac{1}{2}}$ при $a<x<b$, т. е. условие (31) выполняется во всех точках внутри (a, b).

Стеклов рассматривал и такие случаи, когда вес обращается в нуль внутри (a, b), оставаясь неотрицательным. Если вес имеет вид

$$p(x)=(x-a)^{\alpha}(b-x)^{\beta}\prod_{j=1}^{l}(x-x_j)^{2\alpha_j}q(x),$$

где α и $\beta > -1$, α_j — целые положительные числа, а $q(x)$ положительна, ограничена и интегрируема, то, например, для всякой функции $f(x)$, у которой производная $f'(x)$ удовлетворяет условию Липшица (27), доказывается разложимость в ряд Фурье по полиномам Чебышева во всяком замкну-

том промежутке (a_1, b_1), лежащем внутри (a, b) и не содержащем корней $p(x)$. В этом результате малые требования к весу ($q(x)$ только интегрируема) компенсируются большими требованиями к разлагаемой функции (не $f(x)$, а $f'(x)$ должна удовлетворять условию Липшица).

Стеклов установил замкнутость и доказал теоремы разложения и для случая бесконечного промежутка, а именно для промежутка $(0, +\infty)$ при весе $x^{\beta} e^{-x}$ (полиномы Чебышева — Лягерра; $\beta > -1$) и для $(-\infty, +\infty)$ при весе e^{-x^2} (полиномы Чебышева — Эрмита).

5. Вопросы теории замкнутости и разложений по ортогональным функциям явно назрели к концу XIX века, и исследования Стеклова были в мировой литературе первыми работами среди огромного цикла работ в этой области.

Укажем на некоторые моменты из дальнейшей истории этих вопросов, непосредственно связанные с работами Стеклова В 1903 г. Кнезер, улучшив прежний метод Лиувилля, показал возможность разложения по собственным функциям уравнения (4) при тех же условиях на разлагаемую функцию, какая известна для разложения в обычный тригонометрический ряд. Но при этом пришлось наложить тяжелые условия на коэффициенты уравнения. Кнезер пользуется в основном асимптотическим представлением собственных функций $u_k(x)$, т. е. приближенным представлением этих функций при больших значениях k с оценкой порядка ошибки по отношению к величине значка k. Обобщая метод Бонне на разнообразные классы функций и в частности на фундаментальные функции уравнения (4), Стеклов получил в 1907 г. также теоремы разложения с такой же общностью, что и для обычных тригонометрических рядов. Метод построения асимптотических выражений, использованный им в этой работе, называется сейчас *методом Лиувилля — Стеклова*. В дальнейших работах Стеклов не пользуется этим методом, а возвращается к своим прежним методам, связанным большею частью с теорией замкнутости.

Чрезвычайно ярким моментом в истории вопроса было появление в 1904 г. общей теории интегральных уравнений (Гильберт). В эту общую теоретическую схему укладывались все конкретные задачи разложений по фундаментальным функциям. Как это естественно для общей теории, условия разложимости иногда оказывались более грубыми, чем при применении методов, приспособленных к конкретным

случаям. Но в дальнейшем Кнезер и ряд других математиков, усовершенствовав метод интегральных уравнений, получили условия разложимости с большой общностью.

В работе 1905 г. Стеклов отозвался на работы Гильберта по интегральным уравнениям. Применяя тот же метод Шварца — Пуанкаре, которым он обычно пользовался и о котором мы будем говорить ниже, он исследовал решение интегральных уравнений, ядром которых является произведение некоторой положительной функции $p(M)$ на функцию Грина G $(M; N)$. Он показал, что решение есть мероморфная функция параметра и что вычеты этой функции приводят к фундаментальным функциям (ненулевые решения однородного уравнения). Далее он доказал замкнутость системы фундаментальных функций и теорему разложений для функций, представимых через ядро. Функция Грина, как известно, обращается в бесконечность при совпадении точек M и N. Стекловым полностью учтена эта особенность ядра. В дальнейшем он совсем не обращался к теории интегральных уравнений и никогда ею не пользовался.

Вообще его работы по разложениям носили более конкретный характер. Он не был склонен к общим схемам, и, может быть, единственным общетеоретическим моментом в его построениях была созданная им теория замкнутости. Кроме того, существенную роль играл тот метод, с помощью которого Стеклов обычно устанавливал существование фундаментальных функций. Основы этого метода заключаются в работах Шварца (1885 г.) и Пуанкаре (1894 г.), и Стеклов обычно называл этот метод методом Шварца — Пуанкаре. В его руках он стал точным и мощным орудием доказательства существования собственных значений и фундаментальных функций и исследования их свойств. В частности, метод Шварца — Пуанкаре приводил в его работах к доказательству замкнутости системы фундаментальных функций.

Необходимо отметить еще один момент в дальнейшей истории вопроса, а именно проникновение в исследования по теоремам разложений идей и методов современной теории функций вещественного переменного. Я разумею понятия меры множества и измеримых функций, а также интеграл Лебега. Все это сказалось существенным образом и на самих постановках задач. Стеклов оставался в рамках классического анализа. Исходя из уточненного им метода Шварца — Пуанкаре и

созданной им теории замкнутости, он шел самостоятельным путем, стремясь всегда к простоте методов и возможной общности результатов. Он поставил себе проблемы, которые оказались чрезвычайно актуальными и привлекли к себе вскоре внимание выдающихся математиков.

После шестидесятилетнего перерыва, протекшего после работ Лиувилля, Стеклов был одним из первых в деле создания современной нам математической физики. Его непосредственными предшественниками в указанных выше вопросах были только Шварц, Пикар и Пуанкаре.

Таким образом, мной указаны в общих чертах те проблемы, которыми занимался Стеклов в области разложений по ортогональным функциям, и полученные им результаты. Все это я должен был изложить на языке точных математических понятий, что связано с самим характером работ Стеклова, основное содержание которых состоит в доказательствах существования и в выяснении условий приложимости математического аппарата разложений. Но в предыдущем изложении лишь в общих чертах рассмотрены методы, которыми пользовался Стеклов. Отмечено только основное значение теории замкнутости и метода Шварца — Пуанкаре.

В заключении статьи приведем некоторые конкретные методы исследования, которые характерны для его работ. Выберем наиболее существенное и только то, что может быть заключено в рамки небольшой статьи.

6. Начнем с теории замкнутости. Пусть $u_k(x)$ какая-либо ортогональная нормированная система на конечном промежутке (a, b) с весом $p(x)$; обозначим символом $S_n(f)$ среднюю квадратичную погрешность в представлении $f(x)$ отрезком ее ряда Фурье, т. е.

$$S_n(f)=\int_a^b p(x)\,R_n^2(x)\,dx, \tag{33}$$

где $R_n(x)$ определяется формулой (17). Основное значение во всей теории замкнутости играет у Стеклова следующая просто доказываемая формула:

$$\sqrt{S_n(F)}\leqslant\sqrt{S_n(\Phi)}+\sqrt{\int_a^b p(x)\,[F(x)-\Phi(x)]^2\,dx}, \tag{34}$$

где $F(x)$ и $\Phi(x)$ — любые функции. Если $F(x)$ — непрерывная функция, то, согласно теореме Вейерштрасса, можно фиксировать такой полином $\Phi(x)$, что $|F(x)-\Phi(x)|\leqslant\varepsilon$, где ε — любое заданное положительное число. Если предположить, что уравнение замкнутости имеет место для полиномов, то $S_n(\Phi)\to 0$ при $n\to\infty$ и, следовательно, правая часть (34) сколь угодно мала при больших n, откуда следует, что $S_n(F)\to 0$, т. е. уравнение замкнутости имеет место для любой непрерывной функции. Если $F(x)$ только ограничена и интегрируема, то, полагая

$$\Phi(x)=\frac{1}{h}\int_x^{x+h}F(\xi)\,d\xi,$$

выбирая h достаточно малым и опять пользуясь неравенством (34), докажем уравнение замкнутости для $F(x)$. При этом при оценке второго слагаемого правой части (34) надо использовать основное свойство интегрируемых функций. В одной из своих работ Стеклов дал чрезвычайно простое конструктивное доказательство теоремы Вейерштрасса. В другой работе он обходится без теоремы Вейерштрасса в ее общем виде. По существу он очень просто доказывает ее для функций с непрерывными производными до второго порядка, а затем применяет двукратно свой метод сглаживания. В современных терминах из неравенства (34) непосредственно следует, что если уравнение замкнутости имеет место на каком-либо линеале, повсюду плотном в L_2 (семейство измеримых функций таких, что $p(x)[f(x)]^2$ суммируемо), то оно имеет место во всем L_2. Все дело в доказательстве того факта, что линеал полиномов обладает уже этим свойством. Для доказательства этого Стеклов применял свой метод сглаживания. Основными работами, связанными с теорией замкнутости, является его большая работа 1911 г.[1] и две заметки 1916 г. в «Известиях Академии Наук».

7. Переходим к его работе о разложениях по фундаментальным функциям уравнения (9) при предельном условии (10′).

Метод Шварца — Пуанкаре приводит весьма просто к доказательству замкнутости системы фундаментальных функций. Пусть $f(M)=f(x, y, z)$

---

[1] Mémoires de l'Académie des Sciences de St. Pétersbourg, Cl. Ph.-M. VIII S., t. XXX, № 4.

— функция непрерывная с производными до второго порядка в D и равная нулю на S; $G(M;\ N)$ — функция Грина оператора Лапласа для области D; A_k, B_k и C_k — коэффициенты Фурье функций $f(N)$, $\Delta f(N)$ и $G(M;\ N)$ относительно фундаментальных функций $u_k(N)$. Уравнение замкнутости дает

$$\iiint\limits_D G(M;N)\,\Delta f(N)\,d\tau_N = \sum_{k=1}^{\infty} B_k C_k. \tag{35}$$

Основные свойства функции Грина, уравнения для фундаментальных функций $\Delta u_k + \lambda_k u_k = 0$ и интегрирование по частям приводят к равенствам

$$B_k = \iiint\limits_D \Delta f(N)\,u_k(N)\,d\tau = \iiint\limits_D f(N)\,\Delta u_k(N)\,d\tau =$$

$$= -\lambda_k \iiint\limits_D f(N)\,u_k(N)\,d\tau = -\lambda_k A_k,$$

$$C_k = \iiint\limits_D G(M;N)\,u_k(N)\,d\tau_N = \frac{u_k(M)}{\lambda_k},$$

так что (35) дает

$$\iiint\limits_D G(M;N)\,\Delta f(N)\,d\tau_N = -\sum_{k=1}^{\infty} A_k u_k(M). \tag{36}$$

Используя опять свойства функции Грина, легко показать, что левая часть равна $-f(M)$, откуда и следует разложимость $f(M)$ по фундаментальным функциям $u_k(M)$. Отметим, что в работе Стеклова тот факт, что левая часть формулы (35) равна $-f(M)$, доказывается применением уравнения замкнутости к функции

$$u(M) = f(M) + \iiint\limits_D G(M;N)\,\Delta f(N)\,d\tau_N,$$

у которой все коэффициенты Фурье оказываются равными нулю. Абсолютная сходимость ряда

$$\sum_{k=1}^{\infty} A_k u_k(M) = -\sum_{k=1}^{\infty} \frac{B_k}{\lambda_k}\,u_k(M)$$

непосредственно вытекает из неравенства

$$|B_k| \frac{|u_k(M)|}{\lambda_k} \leqslant \frac{1}{2}\left[B_k^2 + \frac{u_k^2(M)}{\lambda_k^2}\right],$$

если принять во внимание, что $u_k(M) : \lambda_k$ суть коэффициенты Фурье функции Грина.

8. Большой цикл работ был посвящен исследованию разложений по фундаментальным функциям уравнения (4) при различных предельных условиях. В этих работах Стеклов обобщал предположения о коэффициентах уравнения и о форме предельных условий. Большая работа по этому вопросу относится к 1901 г.[1] Чрезвычайно широкое обобщение результатов получено в работе 1913 г.[2] Все относящиеся сюда исследования собраны в книге «Основные задачи математической физики», ч. 1 (1922).

Мы начнем с первой из указанных выше работ. Метод Шварца — Пуанкаре, как и выше, дает доказательство замкнутости при предельных условиях (5).

Можно рассматривать $u_k(x)$ как решения «неоднородного» уравнения $u_k''(x) - q(x)u_k(x) = \lambda_k p(x)u_k(x)$, считая правую часть известной функцией от x. Обозначая через $w_1(x)$ и $w_2(x)$ линейно независимые решения уравнения $w''(x) - q(x)\, w(x) = 0$ и применяя к написанному выше уравнению метод вариации произвольных постоянных, получим следующую формулу

$$u_k(x) = \lambda_k \Bigg[M_1(x) \int_a^b p w_2 u_k\, dx + M_2(x) \int_a^b p w_1 u_k\, dx + \\ + \frac{w_1(x)}{\Delta} \int_a^x p w_2 u_k dx - \frac{w_2(x)}{\Delta} \int_a^x p w_1 u_k dx \Bigg], \tag{37}$$

где Δ — постоянная, $M_1(x)$ и $M_2(x)$ — линейные комбинации $w_1(x)$ и $w_2(x)$ с коэффициентами, не зависящими от k. Пользуясь уравнением для $u_k(x)$ и считая, что $f(x)$ имеет производные до второго порядка

[1] Annales de l'Université de Toulouse, t. III.

[2] Mémoires de l'Académie des Sciences de St. Pétersbourg, Cl. Ph.-M., t. XXXI, № 7, 1913.

(причем последняя производная интегрируема) и удовлетворяет предельным условиям, получим, применяя интегрирование по частям, следующее выражение для коэффициентов Фурье $f(x)$:

$$A_k = \int_a^b p f u_k dx = \frac{1}{\lambda_k} \int_a^b q f u_k dx - \frac{1}{\lambda_k} \int_a^b f u_k'' dx =$$

$$= \frac{1}{\lambda_k} \int_a^b (qf - f'') u_k dx,$$

т. е. $A_k = \frac{1}{\lambda_k} C_k$, где C_k — коэффициент Фурье функции $(qf - f'') : p$. Из (37) следует

$$A_k u_k(x) = C_k \left[M_1 \int_a^b p w_2 u_k dx - M_2 \int_a^b p w_1 u_k dx + \right.$$

$$\left. + \frac{w_1(x)}{\Delta} \int_a^x p w_2 u_k dx - \frac{w_2(x)}{\Delta} \int_a^x p w_1 u_k dx \right]. \qquad (38)$$

Но в силу второго основного результата теории замкнутости ряды

$$\sum_{k=1}^{\infty} C_k \int_a^b p w_s u_k dx, \quad \sum_{k=1}^{\infty} C_k \int_a^x p w_s u_k dx \qquad (s=1,\ 2)$$

сходятся, причем сходимость двух последних рядов равномерна. Отсюда, в силу (38), следует равномерная сходимость ряда с общим членом $A_k u_k(x)$ и, опять в силу замкнутости, тот факт, что его сумма равна $f(x)$.

В более позднем мемуаре 1913 г. и упомянутой выше книге для уравнения (4) рассматриваются всевозможные предельные условия, при которых фундаментальные функции ортогональны. В частности, в условиях (5) постоянные h и H уже не считаются положительными. Кроме того, коэффициент $q(x)$ не считается положительным. При этих общих условиях характеристические числа могут быть разных знаков. Пред-

полагается, что разлагаемая функция удовлетворяет условию Липшица (27) или, что равносильно, представляется в виде

$$f(x) = \int_a^x \varphi(\xi)\, d\xi + f(a), \tag{39}$$

где $\varphi(x)$ ограничена и интегрируема. Отметим, что указанная равносильность имеет место, если пользоваться интегралом Лебега, и при этом $\varphi(x)$ почти везде является производной функции $f(x)$, так что, вместо $\varphi(x)$, можно писать $f'(x)$. К функциям, удовлетворяющим условию (39), можно, как известно из теории функций вещественного переменного, применять интегрирование по частям, считая $df(x) = \varphi(x)dx$. Стеклов ссылается при этом обычно на теорему, доказанную Ляпуновым в мемуаре 1904 г. применительно к интегралам Стильтьеса.

Приведем в кратких чертах доказательство одного общего результата, который и служит для доказательства теоремы разложения по фундаментальным функциям уравнения (4) при условии (39). Пусть $u_k(x)$ — любая замкнутая ортогональная нормированная система с положительным весом $p(x)$. Считаем, что существуют ограниченные производные $u_k'(x)$, что, наверное, имеет место для указанных только что фундаментальных функций. При этом $R_n(x)$, определяемая формулой (17), имеет производную

$$R_n'(x) = f'(x) - \sum_{k=1}^{n} A_k u_k'(x). \tag{40}$$

Основной результат состоит в следующем. Если

$$\sigma_n = \int_a^b [R_n'(x)]^2\, dx \leqslant C, \tag{41}$$

где C — постоянная (не зависит от n), то ряд Фурье функции $f(x)$ равномерно сходится. В силу замкнутости отсюда следует, что его сумма равна $f(x)$. Имеем

$$\int_\xi^x \frac{d}{dt}[R_n(t)]^2\, dt = [R_n(x)]^2 - [R_n(\xi)]^2,$$

т. е.

$$[R_n(\xi)]^2=[R_n(x)]^2-2\int_\xi^x R_n(t)\,R_n'(t)\,dt,$$

откуда, применяя неравенство Буняковского — Шварца, получим

$$[R_n(\xi)]^2\leqslant[R_n(x)]^2+2\sqrt{\int_\xi^x[R_n(t)]^2\,dt}\;\sqrt{\int_\xi^x[R_n'(t)]^2\,dt}\leqslant$$

$$\leqslant[R_n(x)]^2+2\sqrt{\int_a^b[R_n(x)]\,dx}\;\sqrt{\int_a^b[R_n'(x)]^2\,dx},$$

или

$$[R_n(\xi)]^2\leqslant[R_n(x)]^2+\frac{2\sqrt{C}}{\sqrt{p_0}}\sqrt{\int_a^b p(x)\,[R_n(x)]^2\,dx}\qquad [p_0 - \text{minimum}\ p(x)].$$

Умножая на $p(x)$ и интегрируя по x на промежутке (a, b), получим

$$Q\,[R_n(\xi)]^2\leqslant\int_a^b p(x)\,[R_n(x)]^2\,dx+\frac{2\sqrt{CQ}}{\sqrt{p_0}}\sqrt{\int_a^b p(x)\,[R_n(x)]^2\,dx},\qquad(42)$$

где

$$Q=\int_a^b p(x)\,dx.$$

В силу замкнутости интеграл, стоящий в правой части (42), стремится к нулю и, следовательно, $R_n(\xi)\to 0$ равномерно для всех ξ из (a, b), что мы и хотели доказать. Таким образом, все дело в проверке условия (41). Это подробно проводится Стекловым при различных предельных условиях. Рассмотрим, например, предельные условия $u(a)= u(b)=0$. При этом функция $f(x)$, кроме (39), должна обращаться в нуль на концах: $f(a)=f(b)=0$. При этом мы имеем, очевидно, $R_n(a)=R_n(b)=0$. Мы можем написать

$$\sigma_n=\int_a^b\Big[f'(x)-\sum_{k=1}^n A_k u_k'(x)\Big]\,R_n'(x)\,dx=\int_a^b f'(x)\,R_n'(x)\,dx-$$

$$-\sum_{k=1}^n A_k\int_a^b u_k'(x)\,R_n'(x)\,dx.$$

Интегрируя последнее равенство по частям, выражая $u_k''(x)$ при помощи уравнения: $u_k''(x) = -\lambda_k p(x)u_k(x) + q(x)u_k(x)$ и принимая во внимание ортогональность $R_n(x)$ и $u_k(x)$ с весом $p(x)$ при $k = 1, 2, \ldots, n$, получим

$$\sigma_n = \int_a^b f'(x)R_n'(x)\,dx + \int_a^b q(x)\,R_n(x)\,f(x)\,dx - \int_a^b q(x)\,[R_n(x)]^2\,dx.$$

Обозначая через q_0 maximum $|q(x)|$ и применяя неравенство Буняковского — Шварца, получим

$$\sigma_n \leqslant \sqrt{\int_a^b [f'(x)]^2\,dx}\,\sqrt{\sigma_n} + \frac{q_0}{\sqrt{P_0}}\sqrt{\int_a^b [f(x)]^2\,dx}\,\sqrt{\int_a^b p(x)\,[R_n(x)]^2 dx} +$$

$$+ \frac{q_0}{P_0}\int_a^b p(x)\,[R_n(x)]^2\,dx.$$

Интегралы, содержащие $p(x)\,[R_n(x)]^2$, стремятся к нулю при $n \to \infty$ и мы приходим к неравенству вида:

$$\sigma_n \leqslant c_1\sqrt{\sigma_n} + c_2,$$

где c_1 и c_2 — постоянные. Из него непосредственно следует, что последовательность чисел σ_n остается ограниченной при возрастании n.

Все указанные выше методы должны во всяком случае ограничиваться областью непрерывных функций $f(x)$. Возможность разложения прерывных функций дает метод асимптотических представлений. В частности, Стеклов строит асимптотические формулы следующего типа:

$$A_k u_k(x) = \varphi(t)\;[B_k \cos(\nu_k t + \tau_k) + \alpha_k w_k],$$

где t — новая переменная, введенная вместо x, ν_k и τ_k — постоянные; ряд, составленный из $|\alpha_k|$, сходится, а $w_k(x)$ и $w_k'(x)$ остаются ограниченными при возрастании k. Далее используется формула

$$\lim_{h\to 0}\frac{1}{4h^2}\int_{x_0-h}^{x_0+h} dx \int_{x-h}^{x+h} f(\xi)\,d\xi = \frac{f(x_0-0)+f(x_0+0)}{2},$$

которая не раз встречается в работах Стеклова; и вопрос сводится к исследованию рядов

$$\sum_{k=1}^{\infty} B_k \cos(\nu_k t + \tau_k) \quad \text{и} \quad \sum_{k=1}^{\infty} B_k \frac{\sin \nu_k h}{\nu_k h} \cos(\nu_k t + \tau_k).$$

9. Переходим теперь к работам последних лет жизни Стеклова, посвященным разложениям функции $f(x)$ по полиномам Чебышева. Пусть

$$\int_a^b p(x) \varphi_k(x) \varphi_l(x) dx = \begin{cases} 0 \text{ при } k \neq l, \\ 1 \text{ при } k = l, \end{cases} \tag{43}$$

где $\varphi_k(x)$ $(k=0, 1, 2, \ldots)$ — полином степени k с положительным старшим коэффициентом, причем промежуток (a, b) считается конечным. Одна из теорем разложения относится к случаю, когда

$$p(x) = (x-a)^{\alpha} (b-x)^{\beta} q(x) \qquad (\alpha \text{ и } \beta > -1), \tag{44}$$

где $q(x)$ положительна и удовлетворяет условию Липшица (27), т. е.

$$q(x) = \int_a^x q'(\xi) d\xi + q(a),$$

и $f(x)$ также удовлетворяет условию Липшица. Исходным пунктом является легко поверяемое тождество

$$p(x)(x-a)(b-x)[R_n(x)]^2 = \int_a^x p(t) \psi(t) [R_n(t)]^2 dt +$$

$$+ \int_a^x p(t) \varphi(t) [R_n(t)]^2 dt + 2 \int_a^x p(t)(t-a)(b-t) R_n(t) R_n'(t) dt, \tag{45}$$

где

$$\varphi(x) = a + b - 2x, \qquad \psi(x) = \alpha(b-x) - \beta(x-a) + (x-a)(b-x) \frac{q'(x)}{q(x)}.$$

Если в первых двух интегралах правой части (45) вынести maximum $|\varphi(x)|$ и $|\psi(x)|$ за знак интеграла и распространить интегрирование на весь промежуток (a, b), то мы получим величину $S_n(f)$, которая стремится к нулю в силу замкнутости. Применяя к третьему интегралу

неравенство Шварца, получим для него оценку в виде корня квадратного из произведения двух интегралов, один из которых

$$\int_a^b p(x)(x-a)(b-x)[R_n(x)]^2\,dx$$

стремится к нулю по тем же соображениям, что и выше. Относительно второго интеграла, имеющего вид

$$\int_a^b p(x)(x-a)(b-x)[R_n'(x)]^2\,dx, \qquad (46)$$

доказывается, что он остается ограниченным при возрастании n (это центральный пункт доказательства). После этого тождество (45) показывает, что $p(x)(x-a)(b-x)[R_n(x)]^2$ равномерно в (a, b) стремится к нулю при $n\to\infty$, а отсюда следует, что ряд Фурье $f(x)$ равномерно сходится и дает $f(x)$ во всяком промежутке (a_1, b_1), лежащем внутри (a, b). При доказательстве ограниченности интегралов (46) используется условие Липшица для $f(x)$.

Переходим теперь к работе о полиномах Чебышева, в которой Стеклов не пользуется теорией замкнутости. Пусть $\psi_k(x)$ — полиномы, которые отличаются от $\varphi_k(x)$ лишь постоянным множителем и в которых старший коэффициент равен единице. При этом имеют место очевидные формулы

$$\varphi_k(x)=\frac{\psi_k(x)}{\sqrt{Q_k}}, \text{ где } Q_k=\int_a^b p(x)[\psi_k(x)]^2\,dx.$$

Исходной в работе служит следующая формула Дарбу для остаточного члена

$$R_n(x)=\frac{1}{Q_n}\int_a^b p(t)F(x, t)[\psi_{n+1}(x)\psi_n(t)-\psi_n(x)\psi_{n+1}(t)]\,dt, \qquad (47)$$

где

$$F(x, t)=\frac{f(x)-f(t)}{x-t}. \qquad (48)$$

Формулу (47) можно переписать в виде

$$R_n(x)=\frac{\sqrt{Q_{n+1}}}{\sqrt{Q_n}}\left[N_n(x)\varphi_{n+1}(x)-N_{n+1}(x)\varphi_n(x)\right], \tag{49}$$

где

$$N_k(x)=\int_a^b p(t)F(x, t)\varphi_k(t)\,dt. \tag{50}$$

Между полиномами $\psi_k(x)$ имеет место известная рекуррентная формула

$$\psi_{n+1}(x)=(x+b_{n+1})\psi_n(x)-c_{n+1}\psi_{n-1}(x),$$

где b_{n+1} и c_{n+1} — постоянные. Умножая эту формулу на $p(x)\psi_{n+1}(x)$, интегрируя по (a, b) и принимая во внимание формулу ортогональности

$$\int_a^b p(x)\psi_k(x)P_{k-1}(x)\,dx=0,$$

где $P_{k-1}(x)$ — любой полином, степень которого $\leqslant k-1$, получим

$$Q_{n+1}=\int_a^b p(x)\,x\psi_{n+1}(x)\psi_n(x)\,dx=\int_a^b p(x)\,(x-a)\psi_{n+1}(x)\psi_n(x)\,dx,$$

и, в силу неравенства Буняковского — Шварца,

$$Q_{n+1}^2\leqslant\int_a^b p(x)\,[\psi_{n+1}(x)]^2\,dx\cdot\int_a^b p(x)\,(x-a)^2\,[\psi_n(x)]^2\,dx,$$

т. е.

$$Q_{n+1}^2\leqslant Q_{n+1}(b-a)^2\,Q_n,$$

откуда

$$Q_{n+1}:Q_n\leqslant(b-a)^2.$$

После этого (49) дает оценку

$$|R_n(x)|\leqslant(b-a)\left[|N_n(x)|\,|\varphi_{n+1}(x)|+|N_{n+1}(x)|\,|\varphi_n(x)|\right],$$

которая приводит к следующему результату. Если в некоторой точке x величины $|\varphi_k(x)|$ остаются ограниченными при $k\to\infty$ и $N_k(x)\to 0$, то в такой точке ряд Фурье по полиномам Чебышева сходится и дает

14*

$f(x)$. В силу (50), $N_k(x)$ суть коэффициенты Фурье функции $F(x, t)$ как функций аргумента t. Полагая

$$F(x, t) = \sum_{k=0}^{n} N_k(x)\varphi_k(t) + R_n(x, t),$$

получим, в силу (18),

$$\int_a^b p(t)\,[R_n(x, t)]^2\,dt = \int_a^b p(t)\,[F(x, t)]^2\,dt - \sum_{k=0}^{n}[N_k(t)]^2 \geqslant 0.$$

При этом считается, что интеграл

$$\int_a^b p(t)\,[F(x, t)]^2\,dt \tag{51}$$

имеет конечное значение. Из последнего неравенства следует, что ряд из $[N_k(x)]^2$ сходится, и потому $N_k(x) \to 0$. Итак, ограниченность $|\varphi_k(x)|$ и существование интеграла (51) гарантируют разложимость $f(x)$. Если $p(x)$ ограниченная функция, то интеграл (51), очевидно, существует, если только $f(x)$ удовлетворяет общему условию Липшица

$$|f(y) - f(x)| \leqslant m\,|y - x|^\gamma \qquad \left(\frac{1}{2} < \gamma \leqslant 1\right). \tag{52}$$

Далее Стеклов показывает, что если не прибегать к приведенному только что рассуждению, а производить непосредственную оценку интеграла (50), то можно доказать теорему разложения в предположении условия (52) при любых $0 < \gamma \leqslant 1$. Остается, конечно, условие ограниченности $|\varphi_k(x)|$. При оценке указанного интеграла путь интегрирования разбивается на три части: $(a, x-\varepsilon)$, $(x-\varepsilon, x+\varepsilon)$, $(x+\varepsilon, b)$, и наряду с $F(x, t)$ создается новая функция $F_1(x, t)$, непрерывная на всем промежутке (a, b), совпадающая с $F(x, t)$ на первой и третьей части и равная полиному 1-й степени от t на второй части. Функция $N_k(x)$ записывается в виде

$$N_k(x) = \int_a^b p(t)F_1(x, t)\varphi_k(t)\,dt + \int_{x-\varepsilon}^{x+\varepsilon} p(t)F(x, t)\varphi_k(t)\,dt -$$
$$- \int_{x-\varepsilon}^{x+\varepsilon} p(t)F_1(x, t)\varphi_k(t)\,dt.$$

Малость первого интеграла получается путем аппроксимации $F_1(x, t)$ полиномами. Малость остальных двух получается за счет ε и ограниченности $p(x)$ в любом интервале, содержащемся внутри (a, b).

В своих работах по разложениям Стеклов не ограничивался только доказательством существования разложения, но он занимался и оценкой величины остаточного члена $R_n(x)$. На этом вопросе мы не останавливались.

В настоящее время имеется огромная литература по теории замкнутости и теоремам разложений. Теория интегральных уравнений объединила прежние методы. Современная метрическая теория функций вещественного переменного дала возможность получить ряд тонких результатов по теоремам разложений. Замкнутые системы дали возможность отождествить в известном смысле функциональное пространство функций с интегрируемым квадратом с пространством последовательностей чисел $(x_1, x_2, \ldots)$, сумма квадратов которых сходится. Тем более мы должны вспомнить того, кто был одним из первых на пути, по которому потом пошли многие.

Академик

И. Г. Петровский

К ТЕОРИИ УРАВНЕНИЙ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

Одна из самых старых и основных задач в теории уравнений с частными производными — это задача Коши. К этой задаче приводят многие физические рассмотрения. Часто, например, бывает, что физическое явление описывается функциями $u_1, \ldots, u_N$, зависящими от времени и пространственных координат x, y, z, причем эти функции удовлетворяют N дифференциальным уравнениям с частными производными второго порядка по t, x, y, z; в начальный момент $t=0$ задаются сами функции u_k и их производные по t (скорость процесса). Задача Коши в этом случае состоит в том, чтобы по этим начальным условиям, заданным при $t=0$, определить дальнейшее течение процесса, т. е. определить функции $u_1, \ldots, u_N$ при $t>0$. В качестве совсем частного примера рассмотрим уравнение колебаний струны

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}.$$

Задача Коши для него состоит в том, чтобы определить его решение $u(t, x)$ при $t>0$, если известны значения самой функции u при $t=0$ (начальная форма струны) и ее производной по t (начальные скорости точек струны). В более общей форме задачу Коши можно сформулировать так.

Пусть дано уравнение вида

$$\frac{\partial^n u}{\partial t^n} = F\left(t,\ x_1, \ldots, x_p,\ u, \ldots, \frac{\partial^k u}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_p^{k_p}} \ldots\right), \qquad (1)$$

где в правую часть входят независимые переменные $t,\ x_1, \ldots, x_p$, неизвестная функция u и ее частные производные по этим независимым переменным до n-го порядка включительно, причем всегда $k_0 < n$. Для простоты записи во всем дальнейшем я буду говорить главным образом об одном уравнении с одною неизвестною функцией, хотя почти все естественно распространяется на системы уравнений. Задача Коши для уравнения (1) в ее наиболее простой формулировке состоит в том, чтобы найти решение этого уравнения, которое при $t=0$ удовлетворяет условиям

$$\frac{\partial^k u}{\partial t^k} = \varphi_k(x_1, \ldots, x_p), \qquad k=0,\ 1, \ldots,\ n-1.$$

Основной является следующая теорема С. В. Ковалевской: если функция F голоморфна по всем своим аргументам вблизи некоторых их значений, которые для простоты мы будем считать нулевыми, и если функции φ_k также голоморфны по всем своим аргументам, причем они обращаются в нуль вместе со всеми своими производными до n-го порядка, когда все $x_k=0$, то задача Коши имеет одно и только одно голоморфное решение в некоторой окрестности начала координат.

Важной вехой в дальнейшем изучении задачи Коши было следующее замечание Адамара [1]. Для уравнения Лапласа

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial x^2} = 0$$

можно как угодно малым изменением начальных значений $u(0,\ x)$ и $u_t'(0,\ x)$ (условий Коши) как угодно сильно изменить решение задачи Коши в какой угодно близости от прямой $t=0$. Для этого надо к прежнему решению $u(t,\ x)$ прибавить

$$v(t,\ x) = \frac{1}{n^k} \sin nx \ \frac{e^{nt} - e^{-nt}}{2},$$

где n и k — положительные постоянные. При достаточно большом n функция v и все ее частные производные по t и x до $(k-1)$-го порядка

как угодно малы при $t=0$, а при $t>\varepsilon>0$ функция $v(t, x)$ принимает как угодно большие значения. Поэтому решение задачи Коши для таких уравнений, как уравнение Лапласа, не представляет физического интереса. В самом деле, начальные условия находятся обычно из опыта измерением, а потому не абсолютно точно. А как угодно малые ошибки, допущенные в определении этих начальных данных, могут как угодно сильно исказить решение задачи Коши и притом в любой близости от линии $t=0$, на которой эти начальные условия задаются. По этой же причине нельзя надеяться даже приближенно получить решение задачи Коши для уравнения Лапласа при неаналитических начальных условиях, аппроксимируя их аналитическими. Ограничиться же только аналитическими начальными условиями невозможно при физических рассмотрениях.

Мы будем говорить, что задача Коши поставлена корректно для уравнения (1) в некоторой области G пространства $(t, x_1, \ldots, x_p)$, примыкающей к куску G_0 гиперплоскости $t=0$, если выполнены следующие условия:

1° Для каждой системы заданных на G_0 начальных функций $\varphi_k(x_1, \ldots, x_p)$, $k=0, 1, \ldots, n-1$, дифференцируемых достаточно большое, но конечное число L раз, существует в области G одно и только одно решение задачи Коши. При этом, если уравнение (1) нелинейно, достаточно ограничиться рассмотрением функций φ_k, отличающихся всюду на G_0 вместе со всеми их частными производными до n-го порядка включительно достаточно мало от некоторых функций $\bar{\varphi}_k(x_1, \ldots, x_p)$ и их соответствующих производных.

2° Для каждого положительного ε существует такое положительное η, что если все заданные на G_0 функции φ_k и их производные до L-го порядка изменить меньше, чем на η, то решение задачи Коши изменится в области G меньше, чем на ε.

Уже давно была доказана корректность постановки задачи Коши для отдельных уравнений. Простейшее из них — «волновое уравнение»

$$\frac{\partial^2 u}{\partial t^2}=\frac{\partial^2 u}{\partial x_1^2}+\ldots+\frac{\partial^2 u}{\partial x_p^2}. \tag{2}$$

Общим классом уравнений, для которых задача Коши поставлена корректно, являются так называемые *гиперболические* уравнения [2, 3].

При определении этого класса я ограничусь для простоты линейными уравнениями вида

$$\frac{\partial^n u}{\partial t^n}=\sum_{k_0+\ldots+k_p\leqslant n} a^{(k_0,\ldots,k_p)}(t, x_1, \ldots, x_p)\frac{\partial^k u}{\partial t^{k_0}\partial x_1^{k_1}\ldots\partial x_p^{k_p}}+f(t, x_1, \ldots, x_p). \quad (3)$$

Будем считать коэффициенты этих уравнений дифференцируемыми достаточно большое конечное число раз.

Рассмотрим в пространстве $(\alpha_1, \ldots, \alpha_p)$ действительную алгебраическую поверхность

$$1-\sum_{k_0+\ldots+k_p=n} a^{(k_0,\ldots,k_p)}(t, x_1, \ldots, x_p)\,\alpha_1^{k_1}\ldots\alpha_p^{k_p}=0. \quad (4)$$

Уравнение (3) называется гиперболическим в некоторой области пространства $(t, x_1, \ldots, x_p)$, если для всех ее точек поверхность (4) в пространстве $(\alpha_1, \ldots, \alpha_p)$ состоит из $\left[\frac{n}{2}\right]$ овалов[1], вложенных последовательно друг в друга и заключающих внутри себя начало координат, а при нечетном n еще из так называемого «непарного куска», гомеоморфного проективной гиперплоскости. Все овалы и непарный кусок не должны иметь общих точек. Тогда каждая действительная прямая в пространстве $(\alpha_1, \ldots, \alpha_p)$, проходящая через начало координат, пересекает поверхность (4) в n действительных и различных точках. Овалы могут быть бесконечными. Чтобы объяснить, что значит тогда, что один овал вложен в другой, будем представлять действительное проективное пространство $(\alpha_1, \ldots, \alpha_p)$, как сферу с отождествленными концами диаметров. Тогда каждому овалу будет соответствовать двуполостный конус, вырезающий на сфере два овала. Мы будем говорить, что первый овал вложен во второй, если конус, соответствующий первому, вложен в конус, соответствующий второму. Непарному куску на сфере соответствует поверхность, гомеоморфная сечению этой сферы гиперплоскостью, проходящей через центр сферы.

[1] $[x]$ означает целую часть от x. Овалом мы называем $(p-1)$-мерную поверхность, гомеоморфную $(p-1)$-мерной сфере; это значит, что между точками $(p-1)$-мерного овала и $(p-1)$-мерной сферы можно установить взаимно однозначное и взаимно непрерывное соответствие.

Если два овала поверхности (4) касаются между собою или если один из овалов касается непарного куска хотя бы в одной точке, то, как показывают примеры, на корректность постановки задачи Коши могут оказывать влияние уже коэффициенты при младших производных. Если коэффициенты уравнения (3) зависят только от t, то требование, чтобы действительная поверхность (4) состояла из $\left[\frac{n}{2}\right]$ хотя бы пересекающихся овалов, заключающих внутри себя начало координат, и еще из непарного куска, если n — нечетно, является необходимым для корректности постановки задачи Коши. В математической физике встречаются уравнения с частными производными, для которых задача Коши поставлена корректно, но для которых некоторые из овалов поверхности (4) имеют общие точки или даже совпадают; такие уравнения встречаются, например, в теории упругости и кристаллооптике. Но обычно решение задачи Коши для таких уравнений бывает легко исследовать, рассматривая эти уравнения как предельные для гиперболических в том смысле, который мы только что определили.

Я ничего не сказал пока о величине того порядка L производных от начальных данных, которая обеспечивает существование решения задачи Коши у гиперболических уравнений. Для получения дважды непрерывно дифференцируемого решения волнового уравнения (2) L должно быть равным $\left[\frac{p}{2}\right]+2$. С. Л. Соболев [4] очень интересными методами, которые будут иметь много других приложений, нашел, что для общего линейного гиперболического уравнения 2-го порядка достаточно считать L равным $\left[\frac{p}{2}\right]+3$, а для нелинейного гиперболического уравнения это число достаточно повысить на 1. У меня [2, 3] для общего нелинейного гиперболического уравнения n-го порядка $L=4p+n+4$, что примерно в 8 раз больше того, что, повидимому, действительно необходимо.

Значение в какой-нибудь точке области G решения u задачи Коши для гиперболического уравнения (1) определяется значениями начальных функций φ_k не на всей области G_0, а на некоторой ее части. Чтобы описать эту часть области G_0, введем понятие о характеристической

поверхности, имеющее важное значение во всей теории уравнений с частными производными. Направление плоскости

$$t - t^* + \alpha_1 (x_1 - x_1^*) + \ldots + \alpha_p (x_p - x_p^*) = 0, \qquad (5)$$

проходящей через точку $P^* (t^*, x_1^*, \ldots, x_p^*)$, если

$$1 - \sum_{k_0 + \ldots + k_p = n} a^{(k_1, \ldots, k_p)} a (t^*, x_1^*, \ldots, x_p^*) \alpha_1^{k_1} \ldots \alpha_p^{k_p} = 0 \qquad (6)$$

называется *характеристическим* для данного уравнения (3) в данной точке P^*. Поверхность в пространстве $(t, x_1, \ldots, x_p)$, у которой всякая касательная гиперплоскость имеет характеристическое направление, называется *характеристической поверхностью* или *характеристикой*. В дальнейшем, говоря о характеристиках, мы всюду будем иметь в виду только действительные характеристики. Характеристическим конусом с вершиной в точке $P (t, x_1, \ldots, x_p)$ называется такая характеристическая поверхность, которой касается в точке P всякая гиперплоскость, имеющая в этой точке характеристическое направление.

Если коэффициенты гиперболического уравнения (3) постоянны, то характеристический конус с вершиной в точке $P^* (t^*, x_1^*, \ldots, x_p^*)$ можно получить, как огибающую семейства плоскостей (5) в пространстве $(t, x_1, \ldots, x_p)$, найденную при условии, что α_k связаны соотношением (6). Будем относить начальные данные к гиперплоскости $t = 0$. Уравнение пересечения этой гиперплоскости с гиперплоскостью (5) будет иметь вид

$$- t^* + \alpha_1 (x_1 - x_1^*) + \ldots + \alpha_p (x_p - x_p^*) = 0,$$

где α_k связаны уравнением (6). Уравнение (6) можно рассматривать, как тангенциальное уравнение пересечения боковой поверхности характеристического конуса с вершиной в точке P^* и гиперплоскости $t = 0$. По определению гиперболичности алгебраическая поверхность (6) n-го порядка в пространстве $(\alpha_1, \ldots, \alpha_p)$ состоит из $\left[\frac{n}{2}\right]$ овалов, вложенных последовательно друг в друга и заключающих внутри себя начало координат; при нечетном n к ним присоединяется непарный кусок. Тот из этих овалов — пусть это будет овал O_{p}, — который вложен во все другие, должен быть обязательно выпуклым; в противном случае

через некоторую точку, лежащую внутри этого овала, можно было бы провести прямую, пересекающую алгебраическую поверхность (6) более чем в n точках. На гиперплоскости $t=0$ овалу O_1 соответствует по закону двойственности также выпуклый овал O_2.

Другим овалам и непарному куску поверхности (6) на гиперплоскости $t=0$ будут соответствовать поверхности, вообще говоря, не выпуклые, с особыми точками, линиями и т. д. Все они будут расположены внутри O_2. Поэтому на гиперплоскости $t=0$ основанием характеристического конуса будет служить конечная область, ограниченная выпуклой поверхностью O_2. Если бы коэффициенты уравнения (6) были переменными, то основанием характеристического конуса все-таки была бы некоторая конечная область, хотя и не обязательно выпуклая. Можно показать, что значение в точке $P^* (t^*, x_1^*, \ldots, x_p^*)$ области G решения задачи Коши для гиперболического уравнения (3) определяется значениями начальных функций только на основании характеристического конуса с вершиной в точке P^*, причем это основание всегда конечно. Поэтому гиперболическими уравнениями описываются такие физические процессы, которые протекают с конечной скоростью во всех направлениях. Это свойство является в некоторой степени характеристическим для гиперболических уравнений.

Можно уточнить характер зависимости решения задачи Коши от значений начальных данных на основании характеристического конуса. Известно, например, что значение в точке $(t, x_1, \ldots, x_p)$ решения задачи Коши для волнового уравнения (2) при нечетном $p > 1$ зависит от значений начальных данных только на периферии основания характеристического конуса с вершиной в точке $(t, x_1, \ldots, x_p)$. При четном же p, а также при $p=1$ величина $u(t, x_1, \ldots, x_p)$ зависит от значений начальных данных на всем основании этого конуса. Напомним, что характеристическими конусами для уравнения (2) служат обычные круглые конусы с осями, параллельными Ot, и с образующими, составляющими с Ot угол в 45°.

Допустим, что начальные значения u и u_t', которые мы будем считать заданными при $t=0$, отличаются от нуля только внутри маленькой области G_ε, около некоторой точки $(0, x_1^0, \ldots, x_p^0)$. Будем следить за значениями u в точках $(t_1, x_1, \ldots, x_p)$ при фиксированных как-нибудь

значениях $x_1, \ldots, x_p$ и при увеличивающемся, начиная от нуля, t. При нечетном $p>1$ величина $u(t, x_1, \ldots, x_p)$ может отличаться от нуля только на небольшом участке рассматриваемой в пространстве $(t, x_1, \ldots, x_p)$ прямой, параллельной оси Ot, именно на том, где расположены вершины характеристических конусов уравнения (2), периферии оснований которых пересекают область G_ε.

Если же $p=1$ или если p четное, то u обязательно равно нулю только в тех точках этой прямой, где расположены вершины характеристических конусов, основания которых не содержат точек области G_ε и которые, очевидно, образуют некоторый отрезок этой прямой, одним из концов которого служит точка $(0, x_1, \ldots, x_p)$. Во всех же других точках этой прямой $u(t, x_1, \ldots, x_p)$, вообще говоря, отлично от нуля. Следовательно, возмущение, произведенное в начальный момент в некоторой малой окрестности точки $(x_1^0, \ldots, x_p^0)$ при нечетном $p>1$ и при $t>0$ отзывается на значениях функции и только в тех точках пространства $(x_1, \ldots, x_p)$, которые лежат около сферы радиуса t с центром в точке $(x_1^0, \ldots, x_p^0)$; таким образом, от возмущения, произведенного в начальный момент в точке $(x_1^0, \ldots, x_p^0)$ возникает сферическая волна с центром в этой точке, имеющая передний и задний фронты. При четном же p и при $p=1$ возмущение, произведенное в начальный момент в окрестности точки $(x_1^0, \ldots, x_p^0)$, отзывается, вообще говоря, на всех точках, лежащих внутри сферы радиуса t с центром в точке $(x_1^0, \ldots, x_p^0)$. Возникает волна, имеющая резкий передний край и размытый задний. Говорят, что в этом случае происходит *диффузия* (размыв) заднего фронта волны; при нечетном $p>1$ диффузии волн для уравнения (2) не бывает.

Еще в начале текущего века Адамар показал, что для линейных гиперболических уравнений второго порядка с переменными коэффициентами при четном p всегда имеет место диффузия волн. Вопрос о диффузии волн для общих линейных гиперболических уравнений второго порядка при нечетном p долгое время оставался открытым. Только в 1939 г. Матисон [5] доказал следующее. При $p=3$ единственным классом линейных гиперболических уравнений второго порядка, у которых нет диффузии волн, являются уравнения, получающиеся из (2) в результате одной из следующих трех операций:

1) умножения обеих частей уравнения на некоторую функцию от $t, x_1, \ldots, x_p$,

2) линейной замены неизвестной функции,

3) перехода от независимых переменных $t, x_1, \ldots, x_p$ к новым независимым переменным.

Методы, которыми пользовался Матисон, принципиально применимы и к гиперболическим уравнениям с любым нечетным числом независимых переменных. Но при $p > 3$ вычисления становятся громоздкими и до конца не доведены.

Я изучал [6] аналогичные вопросы для общих гиперболических систем. Для упрощения записи ограничусь опять при изложении этих результатов случаем одного линейного уравнения (3). Боковая поверхность его характеристического конуса с вершиной в точке $(t^*, x_1^*, \ldots, x_p^*)$ разбивает его основание на плоскости $t = t_0$, вообще говоря, на несколько областей. Одну из этих областей G_{t_0} мы будем называть *лакуной*, если при любых изменениях начальных данных (лишь бы они оставались достаточно гладкими) только внутри G_{t_0} решение задачи Коши для уравнения (3) не изменяется в точке $(t^*, x_1^*, \ldots, x_p^*)$. Для определенности будем считать $t_0 < t^*$. Если t_0 достаточно близко к t^*, то каждой из областей $G_{t_0}^*$, на которые поверхность характеристического конуса с вершиной в точке $(t^*, x_1^*, \ldots, x_p^*)$, построенного для уравнения с постоянными коэффициентами

$$\frac{\partial^n u}{\partial t^n} = \sum_{k_0 + \ldots + k_p = n} a^{(k_0, \ldots, k_p)} (t^*, x_1^*, \ldots, x_p^*) \frac{\partial^n u}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_p^{k_p}}, \quad k_0 < n, \qquad (7)$$

разбивает его основание на плоскости $t = t_0$, будет соответствовать единственная область G_{t_0}, близкая к $G_{t_0}^*$, построенная для уравнения (3). Имеет место следующая теорема.

Если при всех t_0, достаточно близких к t^*, область G_{t_0}, соответствующая области $G_{t_0}^*$, будет лакуной для уравнения (3), *то область $G_{t_0}^*$ будет лакуной для уравнения* (7).

Следствием этой теоремы, по крайней мере при t^*, достаточно близком к t_0, является теорема Адамара о диффузии волн для линейных гиперболических уравнений второго порядка с четным числом пространственных координат x_k, так как для всех линейных гипербо-

лических уравнений второго порядка с постоянными коэффициентами вида (7) всегда имеет место диффузия волн.

Так же легко показать, что при $p=1$ и $n>1$ у уравнения (3) нет лакун. При $p=1$ вопрос о лакунах легко решается для самых общих гиперболических систем. Необходимым и достаточным условием существования лакуны у такой системы является распадение ее на две независимые системы, причем основания характеристических треугольников, которые теперь играют роль характеристических конусов, у этих систем не должны перекрываться.

Все дальнейшее будет относиться к изучению лакун у линейных гиперболических уравнений с постоянными коэффициентами типа (7). Необходимым и достаточным условием отсутствия диффузии волн у таких систем, очевидно, будет наличие лакуны, которую пересекает проходящая через вершину характеристического конуса прямая, параллельная оси Ot. Лакуны, не разрушающиеся при любых достаточно малых изменениях коэффициентов, мы будем называть *устойчивыми*.

Итак, будем рассматривать уравнение (7) с постоянными коэффициентами $a^{(k_0, \ldots, k_p)}$. Чтобы изучить для этого уравнения вопрос о наличии лакун на лежащем в гиперплоскости $t=0$ основании характеристического конуса с вершиной в точке $(t, x_1^*, \ldots, x_p^*)$, достаточно изучить этот вопрос для характеристического конуса K с вершиной в точке $(t, 0, \ldots, 0)$, так как всякая точка $(t, x_1^*, \ldots, x_p^*)$ может быть переведена в эту точку заменой x_k на $x_k+x_k^*$, отчего вид уравнения (7) не изменится. Уравнение

$$1-\sum_{k_0+\ldots+k_p=n} a^{(k_0, \ldots, k_p)} \alpha_1^{k_1} \ldots \alpha_p^{k_p}=0, \qquad k_0<n \tag{8}$$

называется *характеристическим* для уравнения (7). Его можно рассматривать как тангенциальное уравнение поверхности, представляющей собою пересечение боковой поверхности характеристического конуса K с вершиной в точке $(t, 0, \ldots, 0)$, построенного для уравнения (7) и гиперплоскости $t=0$, если уравнение гиперплоскостей в пространстве $(x_1, \ldots, x_p)$ писать в виде

$$-t+\alpha_1 x_1+\ldots+\alpha_p x_p=0. \tag{9}$$

По закону двойственности каждой точке $(x_1, \ldots, x_p)$ в этом пространстве будет соответствовать плоскость в пространстве $(\alpha_1, \ldots, \alpha_p)$. Чтобы решить вопрос о том, принадлежит ли какая-нибудь точка $(x_1, \ldots, x_p)$ устойчивой лакуне в основании конуса K, возьмем в соответствующей плоскости (9) пространства $(\alpha_1, \ldots, \alpha_p)$ какую-нибудь точку A. Проведем через нее всевозможные прямые с действительными коэффициентами, лежащие в этой плоскости, и рассмотрим их точки пересечения с поверхностью (8). Действительные, равно как и недействительные вместе с предельными к ним, точки их пересечения образуют некоторые $(p-2)$-мерные циклы, которые мы будем обозначать соответственно через $C_{\text{reel}}(t)$ и $C_{\text{imag}}(t)$. Цикл $C_{\text{reel}}(t)$, очевидно, не зависит от выбора точки A. Цикл $C_{\text{imag}}(t)$ заменяется ему гомологичным в комплексном пересечении (8) и (9) при изменении A. Если поверхность (8) не имеет особых точек, то справедливы следующие теоремы.

Теорема 1. *При $n \leqslant p$ необходимым и достаточным условием существования устойчивой лакуны уравнения* (7) *является условие, чтобы соответствующие ее точкам циклы* $C_{\text{reel}}(t)$ *в случае нечетного p (соответственно, циклы* $C_{\text{imag}}(t)$ *в случае четного p) были гомологичны нулю в комплексном пересечении* (8) *и* (9).

Имеются примеры таких лакун.

Для решения вопроса о существовании лакун в том случае, когда $n > p$, надо построить циклы $C_{\text{reel}}(\tau)$ в случае нечетного p (соответственно циклы $C_{\text{imag}}(\tau)$ в случае четного p) для каждой из плоскостей

$$-\tau+\alpha_1 x_1+\ldots+\alpha_p x_p=0, \qquad 0 \leqslant \tau \leqslant t. \tag{10}$$

Поверхность, образованную этими циклами, обозначим через S.

Теорема 2. *Необходимым условием существования устойчивой лакуны у уравнения* (7) *в случае $n > p$ является условие, чтобы соответствующие ее точкам циклы* $C_{\text{reel}}(t)$ *при нечетном p (соответственно циклы* $C_{\text{imag}}(t)$ *при четном p) были гомологичны нулю в комплексном пересечении* (8) *и* (9).

Допустим, что это условие выполнено. Натянем пленки на циклы $C_{\text{reel}}(\tau)$ (соответственно $C_{\text{imag}}(\tau)$) в комплексных сечениях поверхности (8) с плоскостями (10) при $\tau=t$ и $\tau=0$. Возможность последнего легко доказать, заметив, что в силу гиперболичности уравнения (7) действи-

тельное пересечение поверхности (8) и гиперплоскости (10) при $\tau=0$ состоит при четном n из $\frac{n}{2}$ овалов, последовательно вложенных друг в друга; при нечетном n к этим овалам присоединяется еще непарный кусок. Указанное натяжение пленок на циклы $C_{\text{reel}}(\tau)$ или $C_{\text{imag}}(\tau)$ при $\tau=0$ и $\tau=t$ можно делать по-разному. Все циклы Σ, образованные при этом поверхностями S и такими пленками, гомологичны между собою в комплексной поверхности (8) при четном p. Если же p нечетное, то при различных способах натягивания пленок на циклы $C_{\text{reel}}(\tau)$ или $C_{\text{imag}}(\tau)$ при $\tau=0$ и $\tau=t$ могут получиться циклы Σ, не гомологичные между собою в комплексной поверхности (8); в таком случае разность их будет алгебраическим циклом, т. е. она будет гомологична в комплексной поверхности (8) некоторой алгебраической поверхности $\frac{p-1}{2}$ комплексных измерений.

Теорема 3. *При выполнении необходимого условия теоремы 2 необходимым и достаточным условием существования устойчивой лакуны уравнения* (7), *если* $n>p$, *является условие, чтобы соответствующие ее точкам циклы* Σ *были гомологичны нулю в комплексной поверхности* (8), *если* p *четно, или же были алгебраическими или гомологичными нулю в поверхности* (8), *если* p *нечетно.*

Заметим, что если цикл $C_{\text{reel}}(t)$ или цикл $C_{\text{imag}}(t)$, соответствующий какой-нибудь внутренней точке одной из областей G_{t_0} в основании характеристического конуса K, гомологичен нулю в комплексном пересечении (8) и (9), то то же имеет место и для всех других точек той же области G_{t_0}. Совершенно так же, если цикл Σ, соответствующий какой-нибудь внутренней точке области G_{t_0}, гомологичен нулю в комплексной поверхности (8), то в этой поверхности будут также гомологичны нулю циклы Σ, соответствующие всем другим точкам области G_{t_0}. Отметим еще, что при выполнении условия теоремы 3 всегда можно так натянуть пленки на циклы $C_{\text{reel}}(\tau)$ при $\tau=0$ и $\tau=t$, чтобы цикл Σ был гомологичен нулю в комплексной поверхности (8).

Наши исследования о лакунах линейных гиперболических уравнений с постоянными коэффициентами опираются на формулы Герглотца [7], дающие решение задачи Коши для уравнений с постоянными коэффициентами вида (7).

Иногда бывает трудно решить вопрос о том, гомологичен ли нулю какой-нибудь из циклов $C_{reel}(t)$ или $C_{imag}(t)$ в комплексном пересечении (8) и (9). Поэтому представляют интерес легко применимые достаточные признаки отсутствия лакун, хотя бы они и не являлись необходимыми. Такие признаки получила недавно А. М. Давыдова. Они состоят в следующем. Рассмотрим какой-нибудь гладкий кусок поверхности S, ограничивающий одну из тех областей G_{t_0}, на которые боковая поверхность характеристического конуса разбивает его основание. Возьмем на этом куске какую-нибудь точку A. Перенесем в нее начало координат. Пусть новая плоскость $x_1=0$ касается поверхности S в точке A. Ось Ox_1 направим внутрь области нормально к плоскости $x_1=0$. Остальные координатные оси Ox_k выберем в плоскости $x_1=0$ так, чтобы уравнение поверхности S вблизи точки A имело вид

$$x_1 = -\sum_{k=2}^{p_1} x_k^2 + \sum_{k=p_1+1}^{p_2} x_k^2, \qquad p_2 \leqslant p.$$

В правой части не выписаны члены высшего порядка относительно $x_2, \ldots, x_p$. Допустим, что для рассматриваемой точки $p_2=p$. Тогда область G_{t_0} заведомо не будет лакуной, если p и p_1 четны или если p и p_1 нечетны.

Все, что мы говорили об одном линейном гиперболическом уравнении с одною неизвестною функцией, естественно, переносится на системы уравнений.

Введем теперь определение гиперболической системы уравнений. Рассмотрим систему линейных уравнений вида

$$\frac{\partial^{n_i} u_i}{\partial t^{n_i}} = \sum_{j=1}^{N} \sum_{k_0+\ldots+k_p \leqslant n} a_{ij}^{(k_0,\ldots,k_p)}(t_1, x_1, \ldots, x_p) \frac{\partial^k u_j}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_p^{k_p}} + f_i(t, x_1, \ldots, x_p); \qquad (11)$$

$$i=1, \ldots, N, \qquad k_0 < n.$$

Составим вспомогательную систему линейных однородных уравнений

$$\frac{\partial^{n_i} u_i}{\partial t^{n_i}} = \sum_{j=1}^{N} \sum_{k_0+\ldots+k_p = n_j} a_{ij}^{(k_0,\ldots,k_p)}(t, x_1, \ldots, x_p) \frac{\partial^{n_j} u_j}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_p^{k_p}}, \qquad (12)$$

$$i=1, \ldots, N, \qquad k_0 < n_j.$$

В правых частях суммирование производится по всем целым неотрицательным $k_0, \ldots, k_p$, сумма которых равна n_j. Будем в уравнениях (12) рассматривать операции $\frac{\partial}{\partial t}$ и $\frac{\partial}{\partial x_k}$ как символическое умножение, а систему (12) как систему линейных алгебраических уравнений относительно $u_1, \ldots, u_N$. Составим детерминант из коэффициентов этой системы. В раскрытом виде его можно переписать так

$$\frac{\partial^n}{\partial t^n} - \sum_{k_0+\ldots+k_p=n} a^{(k_0, \ldots, k_p)}(t, x_1, \ldots, x_p) \frac{\partial^n}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_p^{k_p}},$$

где суммирование распространяется по всем целым неотрицательным $k_0, \ldots, k_p$, сумма которых равна n, причем $n = n_1 + \ldots + n_N$. Уравнение

$$1 - \sum_{k_0+\ldots+k_p=n} a^{(k_0, \ldots, k_p)}(t, x_1, \ldots, x_p)\, \alpha_1^{k_1}, \ldots, \alpha_p^{k_p} = 0 \qquad (13)$$

называется *характеристическим* для системы (11). Система (11) называется *гиперболической* в области G, если во всей этой области гиперболично уравнение

$$\frac{\partial^n u}{\partial t^n} = \sum_{k_0+\ldots+k_p=n} a^{(k_0, \ldots, k_p)}(t, x_1, \ldots, x_p) \frac{\partial^n u}{\partial t^{k_0} \partial_{x_1}^{k_1} \ldots \partial x_p^{k_p}}.$$

Все, что мы говорили об одном гиперболическом уравнении с одною неизвестною функцией, естественно, переносится на гиперболические системы. При этом в сформулированных выше теоремах 1, 2, 3 под n надо понимать наибольшее из чисел $n_1, \ldots, n_N$.

Среди уравнений, имеющих большое значение в математической физике, кроме указанного выше волнового уравнения, лакуны имеются у системы уравнений упругости и у системы уравнений кристаллооптики в трехмерном пространстве, написанных для однородной среды. Каждая из этих систем состоит из трех линейных уравнений второго порядка с постоянными коэффициентами, а соответствующая им действительная поверхность (13) в пространстве $(\alpha_1, \ldots, \alpha_p)$ состоит из конечных овалов, последовательно вложенных друг в друга. Действительным

15*

плоскостям, пересекающим все эти овалы, соответствуют внешности характеристических конусов. Легко показать, что циклы $C_{reel}(t)$, соответствующие этим плоскостям, гомологичны нулю в их комплексном пересечении с поверхностью (13) и потому для этих уравнений часть плоскости $t=0$, лежащая вне основания характеристического конуса, будет лакуной. Это соответствует тому, что значение решения задачи Коши в вершине характеристического конуса зависит только от значений начальных данных на основании этого конуса. Действительным плоскостям в пространстве $(\alpha_1, \ldots, \alpha_p)$, совсем не пересекающим поверхности (13) для уравнений упругости или кристаллооптики, соответствуют лакуны, так как циклы $C_{reel}(t)$ получаются в этом случае пустыми. Действительная поверхность (13) для системы уравнений кристаллооптики имеет особые точки, а два овала этой поверхности в случае уравнений теории упругости даже совпадают. Однако эти уравнения можно рассматривать как предельные для гиперболических в нашем прежнем смысле, когда поверхность (13) не имела особых точек, и, таким образом, перенести и на них наши прежние результаты. Об устойчивости лакун у системы уравнений упругости или у системы уравнений кристаллооптики можно говорить только условно, допуская только такие их изменения, которые не нарушают гиперболичности.

Система нелинейных уравнений

$$\frac{\partial^{n_i} u_i}{\partial t^{n_i}} = F_i\left(t, x_1, \ldots, x_p, u, \ldots, \frac{\partial^k u_j}{\partial t^{k_0}\partial x_1^{k_1}\ldots\partial x_p^{k_p}}, \ldots\right)$$

$$k=k_0+\ldots+k_p\leqslant n, \quad k_0<n_j, \quad i,\ j=1, \ldots, N$$

называется гиперболической в некоторой области G пространства $(t, x_1, \ldots, x_p)$ вблизи функций $\bar{\varphi}_i^{(k)}(x_1, \ldots, x_p)$, $i=1, \ldots, N$, $k=0, \ldots, n_i-1$, если в этой области гиперболична система

$$\frac{\partial^{n_i} u_i}{\partial t^{n_i}} = \sum_{j=1} \sum_{k_0+\ldots+k_p=n_j} a_{ij}^{(k_0, \ldots, k_p)}(t, x_1, \ldots, x_p)\frac{\partial^{n_j} u_j}{\partial t^{k_0}\partial x_1^{k_1}\ldots\partial x_p^{k_p}},$$

$$i=1, \ldots, N, \qquad k_0<n_j,$$

где

$$a_{ij}^{(k_0, \ldots, k_p)} = \frac{\partial F_i}{\partial \left\{ \dfrac{\partial^k u_j}{\partial t^{k_0} \partial x_1^{k_1} \ldots \partial x_p^{k_p}} \right\}},$$

причем правая часть вычисляется при $\dfrac{\partial^k u_j}{\partial t^k} \equiv \varphi_j^{(k)}$.

На нелинейные гиперболические системы легко переносится все то, что мы говорили о линейных гиперболических уравнениях с переменными коэффициентами.

Классом уравнений, во многих отношениях противоположным классу гиперболических уравнений, будут эллиптические уравнения. Система

$$\sum_{j=1}^{N} \sum_{k_0 + \ldots + k_p \leqslant n_j} a_{ij}^{(k_0, \ldots, k_p)} (x_1, \ldots, x_p) \frac{\partial^{n_j} u_j}{\partial x_1^{k_1} \ldots \partial x_p^{k_p}} = f_i(x_1, \ldots, x_p),$$

$$i = 1, \ldots, N,$$

называется *эллиптической* в некоторой области G пространства $(x_1, \ldots, x_p)$, если для каждой точки этой области G детерминант

$$\left| \sum_{k_0 + \ldots + k_p = n_j} a_{ij}^{(k_1, \ldots, k_p)} (x_1, \ldots, x_p) \alpha_1^{k_1} \ldots \alpha_p^{k_p} \right|$$

есть дефинитная форма по $\alpha_1, \ldots, \alpha_p$. Это определение можно распространить на нелинейные уравнения подобно тому, как это мы сделали для гиперболических уравнений. Свойства решений всех эллиптических систем во многом похожи на свойства аналитических функций одного комплексного переменного; действительная и мнимая части этих функций, как известно, удовлетворяют простейшей эллиптической системе — системе Коши-Римана.

В СССР было сделано много фундаментальных исследований по теории эллиптических уравнений. К кругу вопросов, затронутых в настоящей статье, примыкают работы С. Н. Бернштейна [8]. Решив одну из знаменитых проблем Гильберта, С. Н. Бернштейн доказал аналитичность всех достаточно гладких решений эллиптических уравнений второго порядка с двумя независимыми переменными и одною

неизвестною функцией. Эту теорему можно распространить на любые эллиптические системы [9], продолжая заданное действительное решение рассматриваемой эллиптической системы на комплексную область. Это можно сделать, решая задачу, похожую на задачу Коши для гиперболических уравнений, и при том методами, похожими на те, какими мы пользовались при решении задачи Коши для гиперболических уравнений.

ЛИТЕРАТУРА

1. *J. Hadamard.* Le problème de Cauchy. 1932, стр. 40—41.
2. *И. Г. Петровский.* О задаче Коши для системы линейных уравнений с частными производными в области неаналитических функций. Бюлл. Московск. гос. ун-та, серия А, т. I, вып. 7, 1938.
3. *И. Г. Петровский.* О задаче Коши для системы уравнений с частными производными. Матем. сборн., 1937, т. 2 (44): 3, стр. 815—870.
4. *С. Л. Соболев.* К теории нелинейных гиперболических уравнений с частными производными. Матем. сборн., 1939, т. 5(47):1, стр. 71—99.
5. *M. Mathisson.* Le problème de M. Hadamard relatif à la diffusion des ondes. Acta math., t. 71, 1939, № 3—4, p. 249—282.
6. *И. Г. Петровский.* О диффузии волн и лакунах для гиперболических уравнений. Матем. сборн., 1945, т. 17 (59): 3, стр. 289—370.
7. *G. Herglotz.* Ueber die Integration linearer, partieller Differentialgleichungen mit konstanten Koeffizienten. Ber. der Sächischen Akademie, Bd. 78, 1926, S. 93—126, 287—318.
8. *С. Н. Бернштейн.* Sur la nature analytique des solutions des équations aux derivées partielles du second ordre. Math. Ann., Bd. 59, 1904, S. 20—76; Доказательство теоремы Гильберта об аналитическом характере решений эллиптических уравнений без использования нормальных рядов. Усп. матем. наук, 1941, вып. 8, стр. 82—99.
9. *И. Г. Петровский.* Об аналитичности решений систем уравнений с частными производными. Матем. сборн., 1939, т. 5(47):1, стр. 3—70.

Академик
Н. М. Крылов

О НЕКОТОРЫХ НАПРАВЛЕНИЯХ В ОБЛАСТИ ПРИБЛИЖЕННОГО РЕШЕНИЯ ПРОБЛЕМ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

§ 1. Метод конечных разностей. Проблема нахождения приближенных решений дифференциальных уравнений, т. е. уравнений, выражающих связь между двумя бесконечно близкими состояниями физической системы, представилась вниманию геометров уже при самом основании исчисления бесконечно малых. Причиной этого было то обстоятельство, что лишь весьма немногие типы дифференциальных уравнений допускают интегрирование в конечном виде, т. е. решения их выражаются с помощью конечного числа всем хорошо известных элементарных функций.

Так как производные различных порядков получаются, как известно, переходом к пределу в соответствующих разностных отношениях, то естественно, что первые по времени появления приближенные методы интегрирования уравнений были основаны на введении в рассмотрение таких разностных отношений. Так, например, вместо уравнения

$$\frac{dy}{dx}=f(x, y)$$

вводилось соотношение

$$\frac{\Delta y}{\Delta x}=f(x, y),$$

что и привело Леонарда Эйлера к понятию всем хорошо ныне известной «полигональной линии», базируясь на которой Коши указал путь для доказательства существования интеграла дифференциальных систем типа Коши (т. е. для дифференциальных уравнений с граничными условиями на одном конце интервала) и для гораздо более сложных случаев. Это «разностное» направление в области приближенного интегрирования, рассматриваемое с самых разнообразных точек зрения, было предметом множества работ вплоть до настоящего времени; ему же, как известно, следовал и сам Леонард Эйлер в своих исследованиях в области вариационного исчисления.

В связи с этим надлежит отметить, что переход от дискретного к континууму в механических теориях, столь интересовавших классическую математическую физику, может быть совершен двумя различными путями, а именно: 1) в самих разностных уравнениях рассматриваемой проблемы, подставляя вместо них дифференциальные уравнения, которые и рассматривались как уравнения предельной непрерывной системы, и 2) в самих решениях соответствующих уравнений в конечных разностях. Естественным является поэтому вопрос, приводят ли эти два различных перехода к пределу к одному и тому же результату. Интуиция математиков, и прежде всего физиков, дала утвердительный ответ на этот весьма деликатный вопрос, причем тождественность результатов этих двух предельных переходов и составляет сущность так называемого принципа Релея, который является одним из основных в классической математической физике и который был неоднократно в ней применяем в качестве эвристического метода для решения многих проблем как самим Релеем, так и Анри Пуанкаре и другими учеными.

В сравнительно недавнее время «принцип Релея» был обоснован [1, 4] *путем оценки порядка ошибки, совершаемой на n-ом приближении*, причем применением соответственных приемов, а именно введением в рассмотрение разностей высших порядков [4] (так называемый метод высших разностей) может быть повышен, как и следовало ожидать, порядок малости ошибки n-го приближения.

Предметом изучения явились прежде всего дифференциальные системы простейшего типа

$$\frac{d^2y}{dx^2}+\lambda q(x)\,y=0, \qquad y(0)=y(1)=0, \tag{1}$$

хотя методы исследования, о которых идет речь, прилагаются, конечно, и к системам гораздо более сложным.

Ввиду аналогий, существующих между различными физическими явлениями и позволяющих предпринять изучение многих из них помощью одного и того же аналитического аппарата, дифференциальные системы указанных выше типов имеют особо важное значение в весьма большом числе проблем различных прикладных наук, как, например, теория упругости, сопротивление материалов, различные отделы электротехники и т. д.

Это простое и притом далеко не полное перечисление с очевидностью указывает, до какой степени важно приближенное, и притом возможно более точное (т. е. с возможно более точной оценкой ошибки, совершаемой на n-ом приближении) вычисление так называемых сингулярных (особенных) значений параметра $\lambda = \lambda_k$ (а также и сингулярных функций $\varphi_k(x)$, т. е. решений системы (1)), тем более что в физических вопросах с ними связана важная физическая интерпретация, например критическая скорость в теории вращающихся валов, критическая нагрузка в различных вопросах теории упругости и т. д. Представляло поэтому особый интерес рассмотреть систему в конечных разностях

$$\frac{\Delta^2 y_{k,n}(x_{i-1})}{\Delta x^2} + \lambda_k^{(n)} q(x_i)\, y_{k,n}(x_i) = 0, \qquad y_n(0) = y_n(1) = 0,$$

$$q(x) \geqslant 0, \qquad \Delta x = \frac{1}{n}, \qquad i = 1, 2, 3, \ldots, n-1. \tag{2}$$

соответствующую дифференциальной системе (1), и дать оценку разности $\lambda_k - \lambda_k^{(n)}$, что и было выполнено в цитированных выше работах, равно как и повышение порядка малости оценки ошибки n-го приближения путем применения метода «высших разностей», о чем уже говорилось выше. Все исследования в этом направлении, а также и применение «метода разностей» к различным случаям проблемы так называемого «нагруженного минимума» [4] (где для дифференциальных систем типа (1), но с более сложными граничными условиями применяется «метод разностей», исходя из проблемы «связанного» минимума для некоторого, в соответствии с заданными граничными условиями, «нагруженного» интеграла) по самой сути всех трудностей вопроса не имеют еще законченного характера, ибо целью такого рода исследований является, оче-

видно, получение оценки ошибки n-го приближения *не только в отношении порядка ее малости, но по возможности в наименее мажорированной форме.* Это последнее обстоятельство зависит также от весьма тонкой операции, а именно от *соответствующей аппресиации коэффициентов, входящих в выражение оценки ошибки n-го приближения.*

Такого рода снижение мажораций представляет собой иногда особо трудную работу, причем надлежит также иметь в виду, что для полноты исследования в связи с различными случаями, могущими представиться на практике или в теоретических рассмотрениях, придется иметь дело и с так называемыми «локальными» производными высших порядков, на важность рассмотрения которых обращал еще внимание Стильтьес [5] и где высшие разности в пределе принимают вид, связанный с понятием о так называемых «обобщенных» производных, т. е. таких производных, кои существуют, если имеет место производная в обычном смысле, но которые могут существовать и тогда, когда последнее условие не выполнено.

Например, для случая первого разностного отношения $\frac{f(b) - f(a)}{b - a}$ «локальная» производная в точке X существует, как показал Стильтьес, когда $\lim (b - a) = 0$, но при этом $a \leqslant X \leqslant b$. Аналогичные заключения были сделаны Стильтьесом и для «локальных» производных высших порядков, получаемых переходом к пределу в соответствующих интерполяционных формулах.

§ 2. Метод вариационного алгоритма. В вопросах, непосредственно связанных с нахождением решений дифференциальных систем, теоремы, касающиеся существования решений, т. е. так называемые «теоремы существования», играют для математика бесспорно весьма важную роль, в то время как для физика, так и для инженера проблема существования решения отходит часто на второй план, ибо для них *самый факт существования решения бывает иногда непосредственно ясен на основании лишь чисто физических соображений опытного порядка* (как говорил когда-то еще Феликс Клейн: «ток идет, следовательно интеграл существует»), не говоря уже о том, что часто *незначительное изменение в ходе доказательств, как в этом легко убедиться, позволяет конвертировать доказательство сходимости приближенного процесса в доказательство существования решения.*

Принимая это во внимание, становится вполне допустимой и обоснованной та постановка вопроса приближенного интегрирования, которая и была принята в цитированных выше работах (Н. Н. Боголюбова и Н. М. Крылова), а именно: *предполагая установленным существование единственного интеграла данной дифференциальной системы, требуется вычислить его с ошибкой наперед заданной точности путем нахождения мажорированного выражения для этой ошибки по возможности в форме, могущей быть практически использованной.* С вычислительной точки зрения такого рода постановка проблемы приближенного интегрирования представляет, как нам кажется, те преимущества, что в вопросах практических приложений, когда дело идет о самом вычислении решений, не является особо существенным знать, что применяемый процесс приближенного решения будет *сходящимся* и даже, что порядок малости взятого приближения будет достаточно *хорошим.*

Практически *хорошее* приближение *существенно зависит от значений численных коэффициентов, входящих в выражение мажорированной ошибки и влияние их особо сказывается при малых значениях n, которые именно и представляют особый интерес в трудном вопросе практической оценки ошибки n-го приближения.*

В важном вопросе о том, какое значение для n надо взять, чтобы ошибка n-го приближения была менее наперед заданной величины, не следует забывать, что большие значения для n неприемлемы для практических вычислений (впрочем наличие мощного аппарата современных вычислительных машин в значительной мере расширяет возможности в этом отношении), и таким образом получение возможно менее мажорированной ошибки n-го приближения при том или ином процессе приближенного интегрирования связано с трудностью оценки этой ошибки в форме практически приемлемой не для *больших,* но для *малых* значений n.

В вопросах этого рода необходимо различать действительную ошибку, имеющуюся в результате применения данного метода приближенного вычисления, которая теоретически, вообще говоря, остается неизвестной, и мажорированную ошибку, дающую иногда довольно ясное представление о действительной ошибке, как об этом можно заключить из асимптотического выражения, получаемого для оценки ошибок в цитированных выше работах. Надлежит все же заметить, что с мажорированным выра-

жением для ошибок надо действовать с особой осторожностью, в особенности при попытках сравнения различных приемов приближенного интегрирования с точки зрения их практической приложимости, ибо часто другими путями полученное в мажорированной форме выражение для ошибки может привести к заключениям, обратным тем, которые были получены ранее.

После этих предварительных замечаний общего характера, дающих некоторое представление о генезисе тех направлений, которые развивались в цитированных выше работах, обратимся с точки зрения приближенных вычислений к так называемому основному методу математической физики, созданному классическими исследованиями Штурма, Ж. Лиувилля, Шварца, Пикара, Пуанкаре, Фредгольма, В. А. Стеклова, Ст. Заремба и др., где искомое решение — интеграл дифференциальной системы — ищется в форме разложения по так называемым сингулярным (особенным, фундаментальным) функциям, т. е. решениям соответствующей однородной системы и отвечающим сингулярным значениям параметра.

Но так как эти сингулярные функции, равно как и сингулярные значения параметра, вообще говоря, численно не известны, то (не говоря уже о других неудобствах с вычислительной точки зрения метода интегральных уравнений Фредгольма, синтезирующего по сути основной метод математической физики) надлежит в связи с этим вспомнить скептические слова Пуанкаре [6]: «Таким образом, метод Фредгольма, превосходный для строгого доказательства существования решения, не был еще применен для численных вычислений и не представляется пригодным для этого в своей актуальной форме». Интересные исследования, предпринятые ван-ден-Дунгеном [7], привели его к заключению, что вычисление коэффициентов для так называемой трансцендентной Фредгольма «является все более и более неудобным по мере возрастания их индекса, что весьма неприятно, ибо из точного знания этих коэффициентов вытекает почти непосредственно знание, например, пульсаций скручивания».

Хотя некоторые позднейшие исследования [4] и открывают новые пути в этом направлении (например, рассмотрение такого интегрального уравнения, «ядро» которого является приближением к «ядру» данного), тем не менее на основании всего сказанного становится ясным, почему в вопросе нахождения приближенных решений дифференциальных си-

стем пришлось стать на иной путь, и приближенное решение стали искать в форме *конечного* ряда

$$y_n = \sum_{i=1}^{n} a_i \psi_i(x), \tag{3}$$

расположенного по легко *вычисляемым функциям* (например, тригонометрическим или полиномам), удовлетворяющим граничным условиям поставленной физической проблемы. Коэффициенты a_i этих разложений в различных предложенных методах (многие из которых еще до математической обработки их были даны лишь как *приемы для вычислений*) определялись различно. В так называемом методе [4] Релея — Ритца эти коэффициенты определялись из условий стационарности некоторого интеграла, исходя из того, что данное дифференциальное уравнение (в случае его самосопряженности) рассматривалось как уравнение Эйлера, вытекающее из соответствующей вариационной проблемы. Так, например, в случае дифференциальной системы

$$L(y) = f(x), \qquad y_0 = y(1) = 0, \qquad p(x) \geqslant p_1 > 0, \qquad q(x) \leqslant 0, \tag{4}$$

где

$$L(y) = \frac{d}{dx}\left[p(x)\frac{dy}{dx}\right] + q(x)\,y,$$

на основании сказанного придется иметь дело с условиями «стационарности» интеграла

$$I(y) = \int_0^1 [py'^2 - qy^2 + 2fy]\,dx,$$

от варьирования которого и получилась данная система (4), причем в качестве минимизирующих последовательностей берем (3), где $\psi_i(x)$ — некоторая «полная» система легко вычисляемых функций, как об этом уже упоминалось выше. Квадратичная форма, получаемая в результате подстановки (3) в (5), являлась в случаях, рассмотренных самим Ритцем, всегда определенно положительной, благодаря чему для определения неизвестных коэффициентов a_i, исходя из условий стационарности интеграла (5), всегда получалась система линейных уравнений, разрешимых относительно неизвестных a_i.

Очень простым и изящным анализом [4] Ритц устанавливает сходимость минимизирующей последовательности (3) к предельной непрерыв-

ной функции, обладающей производными и удовлетворяющей исходной дифференциальной системе (4). Несмотря на весь интерес работы Ритца, проблема, им затронутая, не могла считаться разрешенной, ибо его анализ неприложим к особо интересным, даже с точки зрения практических приложений, случаям вычисления сингулярных функций и сингулярных значений параметра (для этих случаев Ритц применял метод вариационного алгоритма лишь как *прием для вычисления* без теоретического обоснования).

Таким образом, вопрос по существу оставался открытым и с теоретической точки зрения (сходимость применяемого алгоритма) был исследован в 1918 г. применением метода бесконечных детерминантов. В согласии, однако, с тем, что говорилось уже выше, как должна быть поставлена проблема приближенного интегрирования дифференциальных систем математической физики, надлежало указать различные приемы и методы не только для вычисления мажораций ошибок n-го приближения, но также и для снижения этих мажораций, что и было выполнено в целой серии работ Н. Н. Боголюбовым и Н. М. Крыловым (см., например, цитированную монографию [4], где в конце приведена подробная библиография вопроса (до 1930 г.), содержащая также указания и на более ранние работы Н. М. Крылова в области приближенного интегрирования). Оставляя в стороне вопрос о различных доказательствах сходимости метода Ритца в случаях, им рассмотренных (что может быть сделано весьма кратко [4] надлежащим применением теоремы Арцеля или результатов, относящихся к так называемому «регулярному» случаю вариационного исчисления), обратимся к проблеме приближенного вычисления сингулярных значений параметра. Приведем в качестве примера при $p(x) = 1$ несколько формул, сюда относящихся, замечая, что для практики желательно, по нашему мнению, установление возможно большего количества такого рода мажорационных формул в зависимости от различных ограничительных условий, налагаемых на коэффициенты данной дифференциальной системы и наиболее часто встречающихся на практике:

$$\left|\frac{\lambda_k^{(n)}-\lambda_k}{\lambda_k}\right| < \frac{\lambda_k^{(n)}}{\pi^4 (n+1)^4}\left[|\lambda_k^{(n)}| \max q^2 + \max\left|\frac{q'^2}{q}\right|\right], \qquad (a)$$

$$\left|\frac{\lambda_k^{(n)}-\lambda_k}{\lambda_k}\right| \leqslant \frac{\lambda_k A}{\pi^2 (m+1)^2 - B\,|\lambda_k|}\,, \qquad (б)$$

где[1]

$$A = \max|q| - \min|q|\sqrt{\frac{\max|q|}{\min|q|}}\,,$$

$$B = 2\max|q|, \qquad \lambda_k < \frac{k^2\pi^2}{\min|q|}\,.$$

В той же монографии [4] приведен ряд мажорационных формул для приближенного вычисления сингулярных функций методом вариационного алгоритма.

§ 3. Указания на некоторые другие приемы и методы приближенного решения проблем математической физики (метод наименьших квадратов и наименьших степеней, метод «тронсонов», метод обобщенного коэффициента Фурье и др.). В этих же работах [2, 3] были получены также мажорации ошибок n-го приближения при применении метода наименьших квадратов, который был предложен еще Буссинеском, но лишь как прием для вычисления. Метод этот тоже базируется на одной проблеме нахождения минимума, но не требует, как метод вариационного алгоритма, «самосопряженности» дифференциальной системы; приближенное решение здесь ищется тоже в виде конечного ряда (3), причем искомые коэффициенты определяются для системы (4) из условий стационарности (необходимые условия минимума) интеграла

$$\int_0^1 [L(y)-f]^2\,dx\,.$$

Упомянем также о методе «тронсонов» [2, 4], который можно было бы назвать методом «постоянных коэффициентов» (интервал интегрирования данного дифференциального уравнения делится на n равных частей, и в каждой из них коэффициенты уравнения рассматриваются как

[1] В случае $q(x)=\text{const}$ формула (б) дает, как и следует ожидать, мажорацию ошибки, равную нулю.

постоянные). Метод этот был применяем как прием для вычисления, но в цитированных выше работах для него были установлены различные мажорации ошибок n-го приближения.

В тех же работах (до 1928 г.) Н. Н. Боголюбова и Н. М. Крылова исследован с точки зрения аппресиации ошибки также и метод так называемого «обобщенного коэффициента Фурье», в котором искомые коэффициенты a_i приближенного решения (3) определяются тоже без отношения к какой бы то ни было минимальной проблеме, из такой системы линейных уравнений

$$\int_0^1 [L(y_m) - f] M(\psi_i) \, dx = 0, \qquad i = 1, 2, 3, \ldots, m,$$

которая при тождественности операторов L и M дает метод наименьших квадратов, а при $M(\psi_i) = \psi_i$ как весьма частный случай заключает систему, примененную *лишь как прием для вычислений* (т. е. без доказательства сходимости и аппресиации ошибки n-го приближения) Б. Г. Галеркиным.

Ретроспективный взгляд на все известные до сих пор приближенные методы интегрирования уравнений математической физики (большинство которых, как ранее было уже замечено, может быть конвертировано путем незначительных изменений в приемах доказательств в *самодовлеющие методы математической физики*) делает приемлемым заключение о том, что разные типы дифференциальных систем математической физики (соответственно тем или иным свойствам коэффициентов, входящих в уравнение) допускают более или менее приемлемую мажорацию ошибки n-го приближения в зависимости от применяемого метода приближенного интегрирования, и в этом отношении до известной степени является допустимым утверждение, что каждой дифференциальной системе приличествует для более удобной оценки ошибки тот или иной определенный прием нахождения приближенного решения, т. е. *каждая дифференциальная система имеет, так сказать, свой метод приближенного интегрирования, причем разных методов, равно как и приемов приближенного решения можно создать, как это следует из всего сказанного, достаточно большое количество.*

ЛИТЕРАТУРА

1. *Н. Н. Боголюбов* и *Н. М. Крылов.* О принципе Релея в теории дифференциальных уравнений математической физики и о методе Эйлера в вариационном исчислении. Ann. of Math., 1928.
2. *Н. М. Крылов.* О некоторых недавних изысканиях в области приближенного решения проблем математической физики. Доклад на VI Интерн. матем. конгр. Atti Cong. Inter. Matem., 1928.
3. *Н. М. Крылов.* О приближенном решении проблем математической физики и инженерного дела. Изв. АН СССР, 1930, стр. 1089—1114.
4. *Н. М. Крылов.* Методы приближенного решения проблем математической физики. Монография на франц. яз., стр. 68 (в конце приведены библиографические указания до 1930 г.). Mémorial des Sc. Math., Fasc. XLIX.
5. *Th. Stieltjes.* Oeuvres complètes, vol. I, p. 67—72.
6. *W. Ritz.* Oeuvres complètes avec une Préface de H. Poincaré. Paris, G. Villars, 1911.
7. *Van den Dungen.* Cours de téchnique des vibrations. Bruxelles, 1926.

Академик

А. Н. Колмогоров

СТАТИСТИЧЕСКАЯ ТЕОРИЯ КОЛЕБАНИЙ С НЕПРЕРЫВНЫМ СПЕКТРОМ

1. Предмет, которому посвящен настоящий очерк, может быть охарактеризован в весьма общих терминах, не связанных с какой-либо специальной областью механики или физики. Для простоты мы ограничимся рассмотрением изменения во времени какого-либо конечного числа величин

$$\xi_1(t),\ \xi_2(t),\ \ldots,\ \xi_s(t).$$

Аргумент t (время) будет, естественно, предполагаться действительным, а величины $\xi_r(t)$, вообще говоря, комплексными. Если величины $\xi_r(t)$ изменяются периодически с периодом ω, то процесс их изменения изучается обычно при помощи математического аппарата рядов Фурье

$$\xi_r(t) \sim \sum_n \alpha_r^{(n)} e^{i 2\pi \frac{nt}{\omega}}. \qquad (1)$$

Благодаря этому колебательный процесс изменения величин $\xi_r(t)$ с периодом ω разлагается на гармонические колебания периодов

$$\omega,\ \frac{\omega}{2},\ \frac{\omega}{3},\ \frac{\omega}{4},\ \ldots$$

Естественным обобщением периодических колебаний являются почти-периодические колебания вида

$$\xi_r(t) \sim \sum_n \alpha_r^{(n)} e^{i\lambda_n t}. \qquad (2)$$

Здесь периоды

$$\omega_n = \frac{2\pi}{|\lambda_n|},$$

вообще говоря, несоизмеримы между собой. Естественно, дальнейшим обобщением представляется переход в формуле (2) от сумм по последовательности специально выделенных частот λ_n к интегрированию по непрерывно меняющейся частоте λ

$$\xi_r(t) \sim \int_{-\infty}^{+\infty} \varphi_r(\lambda) e^{i\lambda t} d\lambda, \qquad (3)$$

или в совсем общей форме, объединяющей дискретный случай (2) и непрерывный случай (3), к представлению колебательных процессов в виде интегралов Стильтьеса

$$\xi_r(t) \sim \int_{-\infty}^{+\infty} e^{i\lambda t} d\Phi(\lambda). \qquad (4)$$

Однако, если ограничиться арсеналом средств классического анализа, то на намеченном сейчас пути встретится существенная трудность. Классические интегралы Фурье вида (3) способны представлять лишь функции, которые при $t \to \infty$ стремятся в среднем к константам. Интегралы вида (4) тоже способны представлять лишь функции, которые разлагаются на две компоненты

$$\xi(t) = \eta(t) + \zeta(t),$$

где $\eta(t)$ — почти-периодическая (т. е. имеет дискретный спектр), а $\zeta(t)$ при $t \to \infty$ в среднем стремится к константе. Таким образом, классический аппарат интегралов Фурье для случаев непрерывного спектра приводит лишь к затухающим колебаниям. Так и говорится в учебниках физики, желающих сохранять математическую строгость.[1]

---

[1] См., например, [15].

Тем не менее, идея о том, что и в случае незатухающих колебаний возможны колебания с непрерывным спектром и что их можно изображать интегралами вида (3), культивируется многими, иногда очень крупными, представителями механики и физики. Такого рода исследования, несмотря на отсутствие в них математической строгости, во многих случаях приводят к правильным и полезным результатам.

Предпосылки для строгого математического обоснования теории незатухающих колебаний с непрерывным спектром по существу содержатся в спектральной теории операторов.[1] Математически строгая спектральная теория функций $\xi(t)$ действительного аргумента t, непосредственно ориентированная на математическое обоснование физических представлений о незатухающих колебаниях с непрерывным спектром, была впервые создана Норбертом Винером в 1925 г.[2] Концепция Винера, однако, не приводит ни к представлениям вида (3), ни к представлениям вида (4). Полную ясность в весь рассматриваемый круг вопросов внесла лишь спектральная теория стационарных случайных процессов, предложенная А. Я. Хинчиным в 1934 г. [1]. Она, в частности, приводит к математически строгому и в то же время физически осмысленному обоснованию представления незатухающих статистически стационарных колебаний в виде (4). Это обоснование обладает, повидимому, вполне достаточной для применений общностью. Представление же незатухающих колебаний в виде (3) и в концепции А. Я. Хинчина не получает строгого математического обоснования. Автору этой статьи представляется, что такое положение вещей следует считать окончательным: в случае непрерывного спектра функциям $\Phi_r(\lambda)$, повидимому, никаким разумным способом нельзя приписать определенные производные $\varphi_r(\lambda)$.

Целью настоящего очерка и является рассмотрение некоторых советских и иностранных исследований, развивающих концепцию А. Я. Хинчина.

2. Все очень простые и законченные результаты, излагаемые далее, получаются благодаря радикальному изменению точки зрения: величины

[1] См. изложение этой теории в статьях Плеснера и Рохлина [16].

[2] Систематическое изложение теории Винера можно найти в четвертой главе его книги «The Fourier Integral and certain of its applications» (Cambridge University, Press, 1933); см. также [17].

$\xi_r(t)$ рассматриваются как случайные величины в смысле, принятом в теории вероятностей. Таким образом, основным предметом изучения делается не индивидуальный вполне определенный колебательный процесс, а закон распределения вероятностей в функциональном пространстве различных возможных вариантов течения такого процесса.

Формальное изложение теории может исходить из таких определений:[1]

s-мерным случайным процессом называется совокупность комплексных случайных величин $\xi_r(t)$, заданных для всех действительных t и для $r=1, 2, \ldots, s$.

Предполагается, что математические ожидания

$$M|\xi_r(t)|^2$$

конечны и процесс стохастически непрерывен в том смысле, что

$$M|\xi_r(t+\Delta)-\xi_r(t)|^2 \to 0$$

при $\Delta \to 0$.

Случайный процесс $\{\xi_1(t), \xi_2(t), \ldots, \xi_s(t)\}$ называется *стационарным*, если, каковы бы ни были $t, t_1, t_2, \ldots, t_n$, sn-мерный закон распределения sn случайных величин

$$\begin{array}{llll} \xi_1(t+t_1), & \xi_1(t+t_2), & \ldots, & \xi_1(t+t_n), \\ \xi_2(t+t_1), & \xi_2(t+t_2), & \ldots, & \xi_2(t+t_n), \\ \ldots & \ldots & \ldots & \ldots \\ \xi_s(t+t_1), & \xi_s(t+t_2), & \ldots, & \xi_s(t+t_n) \end{array}$$

не зависит от сдвига t.

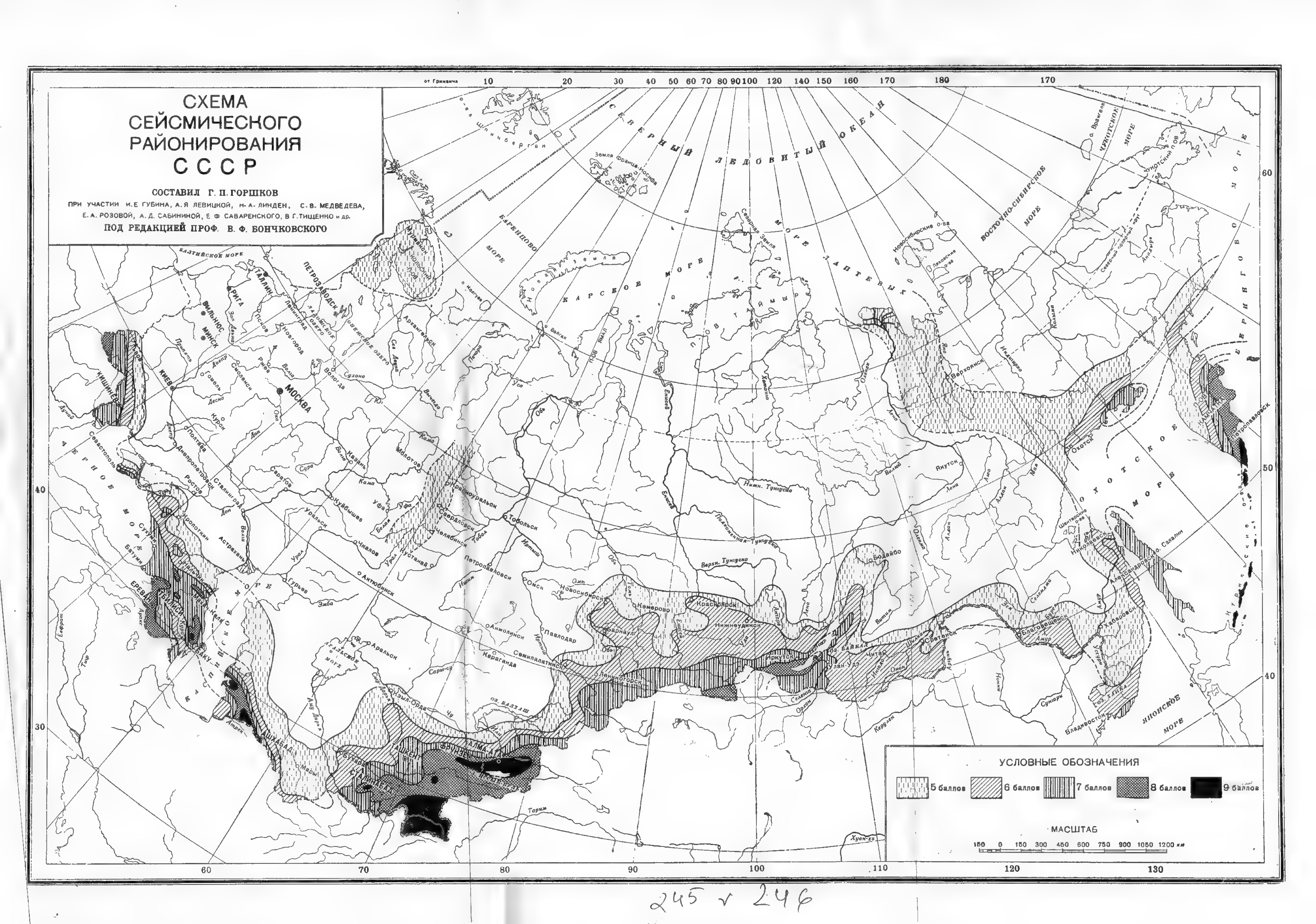
Из стационарности вытекает, в частности, что математические ожидания[2]

$$B_{qr}(\tau)=M\{\xi_q(t+\tau)\,\bar{\xi}_r(t)\} \tag{5}$$

не зависят от t. Само по себе требование независимости от t математических ожиданий (5) называется *стационарностью в широком смысле*.

[1] Употребляемые далее понятия теории вероятностей строго вводятся, например, в моей книге «Основные понятия теории вероятностей». Гостехиздат, 1936.

[2] В равенстве (5) $\bar{\xi}$ обозначает величину, комплексно сопряженную с ξ.


СХЕМА
СЕЙСМИЧЕСКОГО
РАЙОНИРОВАНИЯ
СССР
СОСТАВИЛ Г. П. ГОРШКОВ
ПРИ УЧАСТИИ И. Е ГУБИНА, А. Я ЛЕВИЦКОЙ, Н. А. ЛИНДЕН, С. В. МЕДВЕДЕВА,
Е. А. РОЗОВОЙ, А. Д. САБИНИНОЙ, Е. Ф САВАРЕНСКОГО, В Г. ТИЩЕНКО и др.
ПОД РЕДАКЦИЕЙ ПРОФ. В. Ф. БОНЧКОВСКОГО
УСЛОВНЫЕ ОБОЗНАЧЕНИЯ
5 баллов
6 баллов
7 баллов
8 баллов
9 баллов
МАСШТАБ
150 0 150 300 450 600 750 900 1050 1200 км
СЕВЕРНЫЙ ЛЕДОВИТЫЙ ОКЕАН
БАРЕНЦОВО МОРЕ
КАРСКОЕ МОРЕ
МОРЕ ЛАПТЕВЫХ
ВОСТОЧНО-СИБИРСКОЕ МОРЕ
ЧУКОТСКОЕ МОРЕ
БЕРИНГОВО МОРЕ
ОХОТСКОЕ МОРЕ
ЯПОНСКОЕ МОРЕ
ЧЕРНОЕ МОРЕ
КАСПИЙСКОЕ МОРЕ
МОСКВА
КИЕВ
ТАЛЛИН
РИГА
ВИЛЬНЮС
МИНСК
КИШИНЕВ
ПЕТРОЗАВОДСК
ЕРЕВАН
ТБИЛИСИ
БАКУ
АШХАБАД
ТАШКЕНТ
СТАЛИНАБАД
ФРУНЗЕ
АЛМА-АТА
245 и 246

Все дальнейшее изложение этого и следующего параграфов основано лишь на гипотезе стационарности в широком смысле.

Основной результат А. Я. Хинчина в форме, приданной ему Харальдом Крамером [2], таков:

Функции $B_{qr}(\tau)$ *допускают представление*

$$B_{qr}(\tau)=\int_{-\infty}^{+\infty} e^{i\lambda\tau}dF_{qr}(\lambda), \tag{6}$$

где при любых $\lambda_1<\lambda_2$ *матрица*

$$\| F_{qr}(\Delta_\lambda)\|$$

приращений

$$F_{qr}(\Delta_\lambda)=F_{qr}(\lambda_2)-F_{qr}(\lambda_1)$$

является эрмитовской неотрицательной.

Чтобы понять смысл хинчиновских спектральных функций $F_{qr}(\lambda)$, следует обратиться к дискретному случаю, в котором интегралы (6) превращаются в суммы

$$B_{qr}(\tau)=\sum_n a_{qr}^{(n)} e^{i\lambda_n\tau}. \tag{7}$$

Этот случай был исследован Е. Е. Слуцким [3]. Оказалось, что в предположении (7) имеет место разложение

$$\xi_r(t)\backsim\sum_n \alpha_r^{(n)} e^{i\lambda_n\tau}, \tag{8}$$

где $\alpha_r^{(n)}$ — некоторые случайные величины, однозначно [1] определяемые по заданным $\xi_r(t)$. При этом

$$a_{qr}^{(n)}=M\{\alpha_q^{(n)}\bar{\alpha}_r^{(n)}\}. \tag{9}$$

В частности,

$$a_{rr}^{(n)}=M|\alpha_r|^2. \tag{9'}$$

Естественно было предположить, что и в общем случае

$$F_{qr}(\Delta_\lambda)=M\{\Phi_q(\Delta_\lambda)\bar{\Phi}_r(\Delta_\lambda)\}, \tag{10}$$

$$F_{rr}(\Delta_\lambda)=M|\Phi_r(\Delta_\lambda)|^2, \tag{10'}$$

[1] Однозначность здесь имеет место с точностью до эквивалентных величин т. е. до величин, совпадающих с вероятностью единица.

где $\Phi_r(\Delta_\lambda)$ суть приращения спектральной функции, входящей в представление величин $\xi_r(t)$ вида (4).

Это предположение оказывается правильным. Соответствующее углубление хинчиновской теории стационарных случайных процессов является почти автоматическим следствием ее сведения к спектральной теории однопараметрических групп унитарных операторов, которое указано в работах [4—7] и излагается в следующем параграфе настоящего очерка.

3. Случайные величины $\xi_r(t)$ можно рассматривать как элементы гильбертовского пространства, в котором скалярное произведение определено формулой

$$(\xi, \eta) = M(\xi\bar{\eta}). \tag{11}$$

Легко устанавливается, что любой стационарный процесс $\{\xi_1(t), \xi_2(t), \ldots, \xi_s(t)\}$ в соответствующем гильбертовском пространстве порождает однопараметрическую группу унитарных операторов $\{U^\tau\}$, которые удовлетворяют соотношению

$$U^\tau \xi_r(t) = \xi_r(t+\tau) \tag{12}$$

для всех действительных t и τ и всех $r = 1, 2, \ldots, s$. Операторы U^τ, как известно, допускают представление

$$U^\tau = \int_{-\infty}^{+\infty} e^{i\lambda\tau} dE(\lambda), \tag{13}$$

где $E(\lambda)$ — «разложение единицы». Положив

$$\Phi_r(\lambda) = E(\lambda)\xi_r(0), \tag{14}$$

мы и получим (при надлежащем понимании знака интегрирования) основную формулу (4). Теоретико-вероятностная интерпретация формулы (4) более широко, чем в [6] и [7], изложена в новейших работах Х. Крамера [8], М. Леве [9], А. Бланк-Лапьерра и Р. Фарте ([10] и [11]). Чтобы понять реальный смысл основных спектральных функций $\Phi_r(\lambda)$, естественно рассмотреть их приращения

$$\Phi_r(\Delta_\lambda) = \Phi_r(\lambda_2) - \Phi_r(\lambda_1)$$

и скачки

$$\alpha_r(\lambda) = \Phi_r(\lambda + 0) - \Phi_r(\lambda - 0)$$

в отдельных точках λ. Естественно, что $\Phi_r(\Delta_\lambda)$ и $\alpha_r(\lambda)$, как и $\xi_r(t)$, случайные величины, однако в отличие от $\xi_r(t)$ они не зависят от времени t. Если при каком-либо λ скачок $\alpha_r(\lambda)$ отличен от нуля, то в $\xi_r(t)$ содержится строго-периодическая компонента

$$\xi_r(t, \lambda) = \alpha_r(\lambda) e^{i\lambda t}. \tag{15}$$

В частности, $\alpha_r(0)$ есть среднее (по времени) значение

$$\alpha_r(0) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} \xi_r(t) dt. \tag{16}$$

Аналогичен реальный смысл приращений $\Phi_r(\Delta_\lambda)$ для достаточно коротких отрезков Δ_λ. Именно, компонента

$$\xi_r(t, \Delta_\lambda) = \int_{\Delta_\lambda} e^{i\lambda t} d\Phi(\lambda), \tag{17}$$

соответствующая участку спектра Δ_λ, на любом конечном участке $-T \leqslant t \leqslant +T$ временной оси в случае достаточно короткого Δ_λ сколь угодно хорошо аппроксимируется выражением

$$\Phi_r(\Delta_\lambda) e^{i\lambda t}, \tag{17'}$$

где λ — любая точка из Δ_λ. Естественно, однако, что в случае непрерывного спектра компоненты

$$\xi_r(t, \Delta_\lambda)$$

ни для какого конечного фиксированного интервала Δ_λ не являются строго-периодическими (если только они не равны тождественно нулю). Их временное поведение в случае коротких интервалов Δ_λ аналогично колебаниям маятника со слабым затуханием, возбуждаемого хаотически распределенными малыми случайными толчками.

Компоненты $\xi_r(t, \Delta_\lambda)$ для непересекающихся интервалов Δ'_λ и Δ''_λ между собой не коррелированы, т. е.

$$M\left\{\xi_q(t, \Delta'_\lambda)\, \bar{\xi}_r(t, \Delta''_\lambda)\right\} = 0, \tag{18}$$

если Δ'_λ и Δ''_λ не пересекаются. Для одного интервала Δ_λ

$$M\{\xi_q(t, \Delta_\lambda)\,\overline{\xi_r}(t, \Delta_\lambda)\} = F_{qr}(\Delta_\lambda). \tag{19}$$

В частности,

$$M\,|\xi_r(t, \Delta_\lambda)|^2 = F_{rr}(\Delta_\lambda), \tag{19'}$$

т. е. $F_{rr}(\Delta_\lambda)$ есть не что иное, как среднее значение (по вероятности) квадрата спектральной компоненты $\xi_r(t, \Delta_\lambda)$ величины $\xi_r(t)$.

4. Использование абстрактного аппарата операторов в гильбертовском пространстве могло создать у читателя впечатление, что спектральные компоненты $\xi_r(t, \Delta_\lambda)$ являются какой-то математической фикцией, далекой от возможностей непосредственного эксперимента. В действительности это не так: с любым заданным приближением они могут быть выделены из любого статистически стационарного колебательного процесса при помощи надлежащим образом устроенных *фильтров*. Именно, в принципе с любым желательным приближением можно осуществить прибор, связывающий с любой заданной случайной стационарной функцией $\xi_1(t)$ случайную стационарную функцию

$$\xi_2(t) = \int_0^{+\infty} \xi_1(t-\tau)\, s(\tau)\, d\tau, \tag{20}$$

где $s(\tau)$ подчинено единственному условию

$$\int_0^{\infty} |s(\tau)|^2\, d\tau < +\infty.$$

Если положить

$$s(\tau) = \frac{1}{2\pi i}\,\frac{e^{i\lambda_2(\tau-T)} - e^{i\lambda_1(\tau-T)}}{\tau - T}, \tag{21}$$

то при достаточно большом положительном T

$$\xi_2(t+T) = \frac{1}{2\pi i}\int_{-T}^{+\infty} \xi_1\,(t-\tau)\,\frac{e^{i\lambda_2\tau} - e^{i\lambda_1\tau}}{\tau}\,d\tau$$

будет с вероятностью, сколь угодно близкой к единице, сколь угодно мало отличаться от

$$\xi_1(t, \Delta_\lambda) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \xi_1(t-\tau) \frac{e^{i\lambda_2\tau} - e^{i\lambda_1\tau}}{\tau} d\tau.$$

Заметим еще, что приближенно зависимость (20) можно осуществить в виде

$$L\xi_2 = \xi_1, \tag{22}$$

где

$$L = c_0 + c_1 \frac{d}{dt} + c_2 \frac{d^2}{dt^2} + \ldots + c_n \frac{d^n}{dt^n} \tag{23}$$

и характеристические числа L все имеют отрицательные действительные части. При сделанных предположениях, если $\xi_1(t)$ статистически стационарно, то и $\xi_2(t)$ будет при $t \to +\infty$ стремиться к статистической стационарности. Это предельное статистически стационарное поведение $\xi_2(t)$ определяется формулой

$$\xi_2(t) = \int_{-\infty}^{+\infty} \frac{e^{i\lambda t}}{K(i\lambda)} d\Phi_1(\lambda), \tag{24}$$

$$K(z) = c_0 + c_1 z + \ldots + c_n z^n.$$

Формула (24) имеет самостоятельный интерес независимо от вопроса об экспериментальном определении спектральных компонент $\xi_r(t, \Delta_\lambda)$. Из нее вытекает, что

$$\Phi_2(\lambda) = \int_{-\infty}^{\lambda} \frac{d\Phi_1(\lambda)}{K(i\lambda)}, \tag{25}$$

$$F_{22}(\lambda) = \int_{-\infty}^{\lambda} \frac{dF_{11}(\lambda)}{|K(i\lambda)|^2}. \tag{26}$$

Формулы (24), (25) и (26) содержат в себе обобщение обычной теории резонанса на случай произвольной статистически-стационарной внешней силы. Обычное, не статистическое, изложение этой теории применимо

лишь к случаю дискретного спектра внешней силы, т. е. к случаю $\xi_1(t)$ вида (2). Впрочем, естественно, что без строгого математического обоснования соответствующие рассмотрения для случая непрерывного спектра широко употребляются в физической литературе.

5. В приложениях в случае непрерывного спектра обычно

$$F_{qr}(\lambda) = \int_{-\infty}^{\lambda} f_{qr}(\lambda)\, d\lambda, \tag{27}$$

т. е. формула (6) может быть заменена формулой

$$B_{qr}(\tau) = \int_{-\infty}^{+\infty} e^{i\lambda\tau} f_{qr}(\lambda)\, d\lambda. \tag{28}$$

Этот переход к спектральным плотностям $f_{qr}(\lambda)$ позволяет в той части теории, которая не связана с функциями $\Phi_r(\lambda)$, обходиться без интегралов Стильтьеса.

Самым простым и наиболее важным в применениях является случай, когда законы распределения любого конечного числа величин $\xi_r(t)$ являются *гауссовскими*. В этом случае величины $\Phi_r(\Delta_\lambda)$ тоже подчиняются гауссовским законам распределения. Как раз в этих простейших и типичных для случая непрерывного спектра предположениях (применимость формул (27) и (28) и гауссовских распределений) случайные функции $\Phi_r(\lambda)$ оказываются заведомо недифференцируемыми, и переход от формул (4) к формулам типа (3) представляется безнадежным. На тех участках оси λ, на которых спектральная плотность $f_{rr}(\lambda)$ непрерывна и отлична от нуля, функция $\Phi_r(\lambda)$ имеет тот же характер изменения, какой имеет зависимость от времени координат частицы, подверженной броуновскому движению в случае пренебрежения силами инерции, т. е. для малых $\Delta = \lambda_2 - \lambda_1$ приращения $\Phi_r(\lambda)$ имеют порядок $\sqrt{\Delta}$. Мы имеем здесь новый случай вторжения в математическую физику непрерывных, нигде недифференцируемых функций вейерштрассовского типа. Если в случае броуновского движения более тонкие рассмотрения с учетом сил инерции вновь возвращают нас к дифференцируемым функциям, то здесь недифференцируемость неразрывно связана с самой идеей непрерывного спектра.

За пределами механики и физики колебаний идеи спектрального анализа статистически стационарных процессов находят себе применение (например, в метеорологии) по преимуществу в слегка отличной форме спектральной теории стационарных последовательностей. Этой теории, помимо уже цитированных работ [4 и 5], посвящена книга шведского математика Германа Вольда [12]. Существенные дополнения к теории Вольда были сделаны В. Засухиным [13]. Ясное понимание того обстоятельства, что существование спектра является автоматическим следствием статистической стационарности и вовсе не обязательно указывает на реальное возникновение изучаемого процесса из наложения строго периодических компонент, имело большое значение для критической переоценки так называемой «периодографии», претендовавшей на основное значение в метеорологии и даже экономике. Ставшие классическими в статистической литературе работы в этом направлении принадлежат Е. Е. Слуцкому (особенно [14]).

ЛИТЕРАТУРА

1. *А. Я. Хинчин.* Теория корреляции стационарных стохастических процессов. Усп. матем. наук, 1938, вып. 5, стр. 41—51.
2. *H. Cramer.* Ann. of Math., vol. 41, 1940, p. 215—230.
3. *Е. Е. Слуцкий.* Actualités Scientifiques, № 738, 1938, p. 33—55.
4. *А. Н. Колмогоров.* Изв. АН СССР, серия матем., 1941, № 5, стр. 3—14.
5. *А. Н. Колмогоров.* Бюлл. МГУ (математика), 1941, т. I, вып. 6.
6. *А. Н. Колмогоров.* ДАН СССР, т. 26, 1940, стр. 6—9.
7. *А. Н. Колмогоров.* ДАН СССР, т. 26, 1940, стр. 115—118.
8. *H. Cramer.* Arkiv for Matematik, Astr. och Fysik, vol. 28B, № 12, p. 1—17.
9. *M. Loève.* C. R., vol. 220, 1945, p. 380—382.
10. *A. Blanc-Lapierre* et *R. Fartet.* C. R., vol. 222, 1946, p. 467—468.
11. *A. Blanc-Lapierre* et *R. Fartet.* C. R., vol. 222, 1946, p. 713—714.
12. *H. Wold.* A study in the analysis of stationary time series. Uppsala, 1938.
13. *В. Н. Засухин.* ДАН СССР, т. 33, 1941, стр. 435—437.
14. *Е. Е. Слуцкий.* Сложение случайных причин как источник циклических процессов. Вопросы конъюнктуры, 1927, т. 3, вып. 1, стр. 34—64. Дополненный автором английский перевод в Econometrica, 1937, vol. 5, p. 105—106.
15. *Э. В. Шпольский.* Атомная физика, § 54. Гостехиздат.
16. *А. И. Плеснер* и *В. А. Рохлин.* Усп. матем. наук, 1941, вып. 9; 1946, вып. 11.
17. *N. Wiener.* The harmonic analysis irregular motion. Journ. Math. Phys. Mass. Inst. technology, 1925—26, vol. 5, p. 99—122, 158—191.

ФИЗИКА
*
ГЕОФИЗИКА
*
АСТРОНОМИЯ

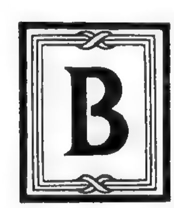
Академик

В. А. Фок

МНОГОЭЛЕКТРОННАЯ ЗАДАЧА КВАНТОВОЙ МЕХАНИКИ И СТРОЕНИЕ АТОМА

§ 1

СВОЙСТВА СИММЕТРИИ ВОЛНОВОЙ ФУНКЦИИ



В квантовой механике определение стационарных состояний системы n электронов сводится к нахождению волновой функции, удовлетворяющей уравнению Шредингера и принципу Паули. Кроме того, во многих задачах можно считать заданным полный спин (собственный момент количества движения) системы электронов; в таком случае волновая функция должна удовлетворять еще одному дополнительному условию.

Волновая функция одного электрона зависит от трех пространственных координат x, y, z и от спиновой переменной σ, принимающей только два значения (например, $\sigma=+1$ и $\sigma=-1$). Обозначив буквой r совокупность трех пространственных координат, мы можем написать волновую функцию одного электрона в виде

$$\psi(x, y, z, \sigma)=\psi(r, \sigma). \tag{1,01}$$

Волновая функция системы n электронов будет зависеть от всех координат и всех спиновых переменных этих электронов. Мы будем иметь

$$\psi=\psi(r_1\sigma_1, r_2\sigma_2, \ldots, r_n\sigma_n). \tag{1,02}$$

Часто бывает удобно обозначать одной буквой x_i совокупность всех переменных, относящихся к электрону номер i (т. е. совокупность его пространственных координат и спиновой переменной). Волновую функцию системы n электронов можно тогда писать в виде

$$\psi = \psi(x_1, x_2, \ldots, x_n). \tag{1,03}$$

Согласно принципу Паули, волновая функция должна быть антисимметрична относительно переменных $x_1, x_2, \ldots, x_n$, т. е. она должна менять знак при перестановке любой пары этих переменных. Например,

$$\psi(x_2, x_1, x_3, \ldots, x_n) = -\psi(x_1, x_2, x_3, \ldots, x_n). \tag{1,04}$$

Переходим теперь к формулировке требования, чтобы система n электронов имела определенный результирующий спин.

В случае одного электрона всякий оператор, действующий на спиновую переменную, может быть представлен как линейная комбинация трех операторов σ_x, σ_y и σ_z, определяемых равенствами

$$\begin{aligned} \sigma_x\psi(r, \sigma) &= \psi(r, -\sigma), \\ \sigma_y\psi(r, \sigma) &= -i\sigma\psi(r, -\sigma), \\ \sigma_z\psi(r, \sigma) &= \sigma\psi(r, \sigma). \end{aligned} \tag{1,05}$$

Нетрудно заметить, что если рассматривать ψ как двухкомпонентную волновую функцию, первая компонента которой равна $\psi(r, +1)$, а вторая $\psi(r, -1)$, то действие операторов σ_x, σ_y, σ_z будет то же, что и действие матриц Паули

$$\sigma_x = \begin{Bmatrix} 0 & 1 \\ 1 & 0 \end{Bmatrix}, \qquad \sigma_y = \begin{Bmatrix} 0 & -i \\ i & 0 \end{Bmatrix}, \qquad \sigma_z = \begin{Bmatrix} 1 & 0 \\ 0 & -1 \end{Bmatrix}. \tag{1,06}$$

Операторы

$$s_x = \frac{1}{2}\sigma_x, \quad s_y = \frac{1}{2}\sigma_y, \quad s_z = \frac{1}{2}\sigma_z \tag{1,07}$$

удовлетворяют перестановочным соотношениям

$$\begin{aligned} s_y s_z - s_z s_y &= i s_x, \\ s_z s_x - s_x s_z &= i s_y, \\ s_x s_y - s_y s_x &= i s_z, \end{aligned} \tag{1,08}$$

которые характеризуют свойства момента количества движения (выраженного в единицах h). Поэтому их можно толковать как операторы для составляющих собственного момента количества движения электрона. В случае многих электронов можно определить аналогично (1,05) операторы σ_{lx}, σ_{ly}, σ_{lz}, действующие на спиновую переменную электрона номер l. Мы будем иметь

$$\begin{aligned} \sigma_{lx}\psi &= \psi(r_1\sigma_1, \ldots, r_l, -\sigma_l, \ldots), \\ \sigma_{ly}\psi &= -i\sigma_l\psi(r_1\sigma_1, \ldots, r_l, -\sigma_l, \ldots), \\ \sigma_{lz}\psi &= \sigma_i\psi(r_1\sigma_1, \ldots, r_l\sigma_l, \ldots). \end{aligned} \tag{1,09}$$

Операторы для составляющих спинового момента количества движения системы электронов, выраженные в единицах h, могут быть определены аналогично (1,07), а именно

$$\begin{aligned} s_x &= \frac{1}{2}(\sigma_{1x} + \sigma_{2x} + \ldots + \sigma_{nx}), \\ s_y &= \frac{1}{2}(\sigma_{1y} + \sigma_{2y} + \ldots + \sigma_{ny}), \\ s_z &= \frac{1}{2}(\sigma_{1z} + \sigma_{2z} + \ldots + \sigma_{nz}). \end{aligned} \tag{1,10}$$

Эти операторы удовлетворяют тем же перестановочным соотношениям (1,08). Из перестановочных соотношений (1,08) можно вывести, что оператор

$$\mathbf{s}^2 = s_x^2 + s_y^2 + s_z^2 \tag{1,11}$$

для квадрата собственного момента количества движения электронов будет коммутировать с каждым из операторов s_x, s_y, s_z и что его собственные значения будут равны $s(s+1)$, где s есть половина целого неотрицательного числа. Если число n электронов — четное, то s есть целое положительное число или нуль. Если же n — нечетное, то s есть полуцелое число (половина нечетного целого числа). В обоих случаях число $\frac{n}{2} - s = k$ будет неотрицательным целым числом. Это число k можно толковать как число пар электронов с компенсированным спином.

При заданном s собственные значения каждого из операторов s_x, s_y, s_z пробегают ряд чисел

$$-s, -s+1, \ldots, s-1, s, \tag{1,12}$$

т. е. всего $2s+1$ значение.

Оператор (1,11) для $\mathbf{s}^2$ может быть представлен в виде

$$\mathbf{s}^2 = n - \frac{n^2}{4} + \sum_{i<j} P_{ij}, \tag{1,13}$$

где символ P_{ij} означает перестановку спиновых переменных σ_i и σ_j.

Требование, чтобы система n электронов имела определенный результирующий спин, может быть теперь записано в виде уравнения

$$\mathbf{s}^2\psi = s(s+1)\psi. \tag{1,14}$$

Построим функцию, удовлетворяющую этому уравнению. Положим $k = \frac{n}{2} - s$ и пусть $\alpha_1, \alpha_2, \ldots, \alpha_k$ есть совокупность k различных чисел, взятых из ряда $1, 2, \ldots, n$. Пусть, далее,

$$F_{\alpha_1\alpha_2\ldots\alpha_k} = F(\sigma_{\alpha_1}\sigma_{\alpha_2}\ldots\sigma_{\alpha_k} | \sigma_{\alpha_{k+1}}\ldots\sigma_{\alpha_n}) \tag{1,15}$$

есть функция от спиновых переменных, симметричная как относительно аргументов $\sigma_{\alpha_1}\sigma_{\alpha_2}\ldots\sigma_{\alpha_k}$, стоящих до вертикальной черты, так и относительно аргументов $\sigma_{\alpha_{k+1}}\ldots\sigma_{\alpha_n}$, стоящих после черты. Введем совокупность $\binom{n}{k}$ функций

$$\psi_{\alpha_1\alpha_2\ldots\alpha_k} = \psi_{\alpha_1\alpha_2\ldots\alpha_k}(r_1r_2\ldots r_n) \tag{1,16}$$

от координат всех электронов (эти функции уже не содержат спиновых переменных). Функции (1,16), так же как и функции (1,15), симметричны относительно значков $\alpha_1, \alpha_2, \ldots, \alpha_k$. Составим теперь сумму

$$\psi(r_1\sigma_1, r_2\sigma_2, \ldots, r_n\sigma_n) = \\ = \sum_{(\alpha_1\ldots\alpha_k)} \psi_{\alpha_1\ldots\alpha_k}(r_1\ldots r_n) F_{\alpha_1\ldots\alpha_k}(\sigma_1\ldots\sigma_n). \tag{1,17}$$

На основании формулы (1,13) можно показать, что функция (1,17) будет удовлетворять уравнению (1,14) со значением s, равным $\frac{n}{2} - k$.

если только координатные функции (1,16) связаны соотношениями

$$\sum_{\alpha} \psi_{\alpha\alpha_2 \ldots \alpha_k} = 0, \tag{1,18}$$

где значок α пробегает все значения из ряда 1, 2, ... n, за исключением значений $\alpha_2, \ldots, \alpha_k$. Число соотношений (1,18) равно $\binom{n}{k-1}$.

Для того чтобы полученная собственная функция оператора $\mathbf{s}^2$ представляла физически возможное состояние системы электронов с заданным спином, необходимо, чтобы она удовлетворяла также принципу Паули, т. е. была антисимметрична относительно переменных x_i. Этого можно достигнуть, выразив по формуле

$$\psi_{\alpha_1\alpha_2 \ldots \alpha_k}(r_1 r_2 \ldots r_n) = \varepsilon(P)\,\psi(r_{\alpha_1} r_{\alpha_2} \ldots r_{\alpha_k} \,|\, r_{\alpha_{k+1}} \ldots r_{\alpha_n}) \tag{1,19}$$

все функции (1,16) через одну функцию

$$\psi = \psi(r_1 r_2 \ldots r_k \,|\, r_{k+1} r_{k+2} \ldots r_n) \tag{1,20}$$

от координат электронов. В формуле (1,19) через $\alpha_1, \alpha_2, \ldots, \alpha_n$ обозначены числа 1, 2, . . . , n, взятые в произвольном порядке, и через P — та перестановка,

$$P = \begin{pmatrix} 1 & 2 & \ldots & n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}, \tag{1,21}$$

которая переводит 1 в α_1, 2 в α_2 и т. д. Символом $\varepsilon(P)$ обозначено число, равное $+1$, если перестановка P четная, и равное -1, если она нечетная.

Функция (1,20) должна удовлетворять следующим условиям симметрии:

1) ψ антисимметрична относительно первых k аргументов [стоящих в (1,22) слева от черты], например:

$$\psi(r_2 r_1 r_3 \ldots r_k \,|\, r_{k+1} \ldots r_n) = -\psi(r_1 r_2 r_3 \ldots r_k \,|\, r_{k+1} \ldots r_n); \tag{1,22}$$

2) ψ антисимметрична относительно последних $n-k$ аргументов (стоящих справа от черты), например:

$$\psi(r_1 r_2 \ldots r_k \,|\, r_{k+2} r_{k+1} r_{k+3} \ldots r_n) = \\ = -\psi(r_1 r_2 \ldots r_k \,|\, r_{k+1} r_{k+2} r_{k+3} \ldots r_n) \tag{1,23}$$

3) ψ обладает свойством циклической симметрии, которое выражается равенством

$$\psi(r_1 \ldots r_{k-1} r_k | r_{k+1} \ldots r_n) = \psi(r_1 \ldots r_{k-1} r_{k+1} | r_k r_{k+2} \ldots r_n) + \\ + \ldots + \psi(r_1 \ldots r_{k-1} r_{k+l} | r_{k+1} \ldots r_{k+l-1} r_k r_{k+l+1} \ldots r_n) + \\ + \ldots + \psi(r_1 \ldots r_{k-1} r_n | r_{k+1} \ldots r_{n-1} r_k). \qquad (1,24)$$

Правая часть этого равенства состоит из $n-k$ членов, в которых аргумент r_k ставится последовательно на место каждого из $n-k$ аргументов справа от черты.

Свойство циклической симметрии является следствием соотношений (1,18). Если обозначить через P циклическую перестановку аргументов $(r_k r_{k+1} \ldots r_n)$, т. е. такую перестановку, в которой каждый член цикла $(r_k r_{k+1} \ldots r_n)$ заменяется следующим членом, а последний член — первым членом, то равенство (1,24) может быть записано в виде

$$(1 + P + P^2 + \ldots + P^{n-k})\psi = 0 \qquad (1,25)$$

при $n-k$ чётном и в виде

$$(1 - P + P^2 - \ldots - P^{n-k})\psi = 0 \qquad (1,26)$$

при $n-k$ нечетном.

В частном случае двух электронов состояние с нулевым спином ($n=2$, $k=1$) описывается симметричной функцией от координат, а состояние со спином $s=1$ ($n=2$, $k=0$) антисимметричной функцией.

Весьма важным примером функции от n аргументов $r_1 r_2 \ldots r_n$, удовлетворяющей трем формулированным выше условиям симметрии, является произведение двух определителей

$$\psi = \Psi^{(1)} \cdot \Psi^{(2)}, \qquad (1,27)$$

где

$$\Psi^{(1)} = \begin{vmatrix} \psi_1(r_1) & \ldots & \psi_1(r_k) \\ \ldots & \ldots & \ldots \\ \psi_k(r_1) & \ldots & \psi_k(r_k) \end{vmatrix}, \quad \Psi^{(2)} = \begin{vmatrix} \psi_1(r_{k+1}) & \ldots & \psi_1(r_n) \\ \ldots & \ldots & \ldots \\ \psi_{n-k}(r_{k+1}) & \ldots & \psi_{n-k}(r_n) \end{vmatrix}. \qquad (1,28)$$

Эти определители составлены из функций

$$\psi_1(r), \ \psi_2(r), \ \ldots, \ \psi_{n-k}(r), \qquad (1,29)$$

зависящих от координат одного электрона, причем больший определитель содержит все $n-k$ функций (1,29), а меньший — только первые k из них.

Предыдущие рассуждения позволили нам выразить волновую функцию (1,02), которая зависит, кроме координат, еще от всех спиновых переменных, через шредингеровскую волновую функцию (1,20), зависящую только от одних координат. При этом как принцип Паули, так и уравнение (1,14) для спинового момента количества движения были учтены вполне строго.

Несмотря на то, что шредингеровская волновая функция не зависит от спиновых переменных, значение результирующего спина отражается на ее свойствах, так как от него зависит ее симметрия.

Этим объясняется тот факт, который на первый взгляд кажется парадоксом: ни уравнение Шредингера, ни волновая функция спиновых переменных не содержит, между тем уровни энергии зависят от значения результирующего спина. Парадокс разъясняется тем, что на волновую функцию, соответствующую уровню с данным спином, дополнительно накладываются условия симметрии, различные для различных значений результирующего спина.

§ 2

ОПЕРАТОР ЭНЕРГИИ И ЕГО СИММЕТРИЯ

Волновая функция, описывающая стационарное состояние многоэлектронной системы, должна быть собственной функцией оператора энергии, который имеет вид

$$H = -\frac{h^2}{2m}\sum_{k=1}^{n}\Delta_k + \sum_{k=1}^{n} U(x_k y_k z_k) + \sum_{k>l=1}^{n}\frac{e^2}{\sqrt{(x_k-x_l)^2+(y_k-y_l)^2+(z_k-z_l)^2}}. \qquad (2,01)$$

Здесь Δ_k есть оператор Лапласа, действующий на координаты электрона номер k, а $U(x, y, z)$ — потенциальная энергия поля, внешнего по отношению к электронам (например, поля ядра для атома или несколь-

ких ядер для молекулы), двойная сумма есть взаимная потенциальная энергия электронов. Постоянные e, m, h имеют обычное значение (заряд электрона, масса электрона, деленная на 2π постоянная Планка). Оператор энергии (2,01) соответствует тому случаю, когда магнитное поле отсутствует; если бы система электронов находилась во внешнем магнитном поле, то оператор энергии содержал бы дополнительные члены, зависящие также и от спина.

Уровни энергии и стационарные состояния системы определяются из уравнения

$$H\psi = E\psi, \tag{2,02}$$

где H есть оператор (2,01). Мы уже указывали на то, что хотя оператор H не содержит спиновых переменных, тем не менее уровни энергии E зависят от квантового числа s (спинового момента количества движения). Объяснение этого обстоятельства заключается, как мы знаем, в том, что от значения s зависят свойства симметрии шредингеровской волновой функции ψ.

В случае атома оператор H обладает сферической симметрией, т. е. его вид не меняется при любом повороте координатных осей в пространстве. Тогда можно подчинить шредингеровскую (координатную) волновую функцию требованию, чтобы она была собственной функцией оператора для квадрата орбитальнсго момента количества движения (квантовое число l) и для составляющей его по одной из осей (квантовое число m). Если эта функция, кроме того, обладает свойствами симметрии, соответствующими определенному значению s, то при помощи нее можно построить волновую функцию со спином [вида (1,02)], которая будет удовлетворять принципу Паули и будет собственной функцией следующих пяти операторов: 1) оператора энергии, 2) квадрата орбитального момента количества движения, 3) квадрата спинового момента количества двпжения, 4) квадрата полного (орбитального плюс спинового) момента количества движения, 5) составляющей полного момента количества движения по одной из осей. Построение достигается при помощи векторной модели [10]; мы не будем на этом останавливаться.

В случае двухатомной молекулы оператор энергии обладает не сферической, а лишь аксиальной симметрией (т. е. не меняется при по-

вороте координатной системы вокруг оси, соединяющей оба ядра). Аксиальная симметрия также может быть использована для введения квантовых чисел и для частичного определения волновых функций.

Использование сферической или аксиальной симметрии системы позволяет вводить квантовые числа и тем самым классифицировать уровни энергии. Однако соображений симметрии недостаточно для определения самих уровней и стационарных состояний. Точное решение уравнения (2,02) представляет (за исключением случая одного электрона) непреодолимые математические трудности. Поэтому приобретает большое значение развитие приближенных методов. Наиболее важным из этих методов является метод согласованного поля, к изложению которого мы и переходим.

§ 3

МЕТОД СОГЛАСОВАННОГО ПОЛЯ

Уравнение для собственных функций оператора энергии может быть получено из вариационного начала

$$\delta W = 0, \tag{3,01}$$

где W есть выражение

$$W = \frac{1}{N}\int \overline{\psi} H \psi dV, \qquad N = \int \overline{\psi}\psi dV. \tag{3,02}$$

В этой формуле мы можем разуметь под ψ шредингеровскую координатную функцию (не зависящую от спиновых переменных). Элементом объема dV конфигурационного пространства будет тогда произведение дифференциалов координат всех электронов

$$dV = dx_1 dy_1 dz_1 \ldots dx_n dy_n dz_n. \tag{3,03}$$

Нормировочный интеграл N мы можем считать заданной постоянной. Физический смысл величины W есть математическое ожидание энергии системы в состоянии ψ. Экстремальное значение W есть уровень энергии E. Чтобы получить наинизший уровень, соответствующий данному значению квантового числа s, мы должны при варьировании интеграла допускать к сравнению все функции ψ, обладающие нужными свой-

ствами симметрии и удовлетворяющие некоторым общим условиям (существование производных, сходимость интегралов). Чтобы получить последующие уровни, мы должны, сверх того, потребовать, чтобы волновая функция была ортогональна ко всем функциям, соответствующим более низким уровням.

С целью упростить решение задачи мы можем наложить на волновую функцию некоторые дополнительные условия, например потребовать, чтобы она выражалась, согласно (1,27) и (1,28), в виде произведения двух определителей. В таком случае, вместо наинизшего уровня, мы получим несколько более высокое значение энергии, которое, однако, будет мало от него отличаться. Подобным же образом мы получим для следующих уровней близкие к ним значения.

Вычислим результат подстановки в W произведения определителей (1,28), причем будем предполагать функции $\psi_p(r)$ ортогональными между собой

$$\int \overline{\psi}_p(r)\,\psi_q(r)\,d\tau = \delta_{pq}, \qquad d\tau = dx\,dy\,dz, \tag{3,04}$$

что, очевидно, не нарушает общности. Для этого представим оператор энергии (2,01) в виде

$$H(r_1, \ldots, r_n) = \sum_{p=1}^{n} (H(\mathbf{r}_p) + \sum_{p>q=1}^{n} \frac{e^2}{|\mathbf{r}_p - \mathbf{r}_q|}, \tag{3,05}$$

где

$$H(\mathbf{r}) = -\frac{h^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + U(x, y, z). \tag{3,06}$$

Мы получим тогда

$$\begin{aligned} W = {} & \sum_{p=1}^{k} \int \overline{\psi}_p(r)\,H(r)\,\psi_p(r)\,d\tau + \sum_{p=1}^{n-k} \int \overline{\psi}_p(r)\,H(r)\,\psi_p(r)\,d\tau + \\ & + \frac{e^2}{2} \iint \frac{\rho^{(1)}(r, r)\,\rho^{(1)}(r', r') - |\rho^{(1)}(r, r')|^2}{|\mathbf{r} - \mathbf{r}'|}\,d\tau\,d\tau' + \\ & + \frac{e^2}{2} \iint \frac{\rho^{(2)}(r, r)\,\rho^{(2)}(r', r') - |\rho^{(2)}(r, r')|^2}{|\mathbf{r} - \mathbf{r}'|}\,d\tau\,d\tau' + \\ & + e^2 \iint \frac{\rho^{(1)}(r, r)\,\rho^{(2)}(r', r')}{|\mathbf{r} - \mathbf{r}'|}\,d\tau\,d\tau'. \end{aligned} \tag{3,07}$$

В этой формуле мы обозначили через $\rho^{(1)}$ и $\rho^{(2)}$ следующие выражения:

$$\rho^{(1)}(r,\ r') = \sum_{p=1}^{k} \overline{\psi}_p(r)\,\psi_p(r'), \tag{3,08}$$

$$\rho^{(2)}(r,\ r') = \sum_{p=1}^{n-k} \overline{\psi}_p(r)\,\psi_p(r'). \tag{3,09}$$

Полученные формулы допускают наглядное толкование. Прежде всего, выражение полной волновой функции системы через волновые функции $\psi_p(r)$ соответствует предположению, что мы можем каждому электрону системы приписать свою волновую функцию (мы можем условно сказать: свою орбиту). При этом электроны распадаются на два «роя»: к первому рою относятся электроны на «орбитах» $\psi_1, \psi_2, \ldots, \psi_k$, а ко второму — электроны на орбитах $\psi_1, \psi_2, \ldots, \psi_{n-k}$. Оба роя отличаются друг от друга противоположным спином. На орбитах $\psi_1, \psi_2, \ldots, \psi_k$ находятся по два электрона с разным спином, а на остальных орбитах $\psi_{k+1}, \ldots, \psi_{n-k}$ по одному электрону с одинаковым спином. Спины электронов на первых k орбитах взаимно уничтожаются, а спины электронов на остальных $n-2k$ орбитах складываются; так как абсолютная величина спина каждого электрона равна $\frac{1}{2}$, то полный спин системы получается равным $\frac{n}{2} - k = s$, как и следовало ожидать. Умноженная на заряд электрона e величина $\rho^{(1)}(r,\ r)$, равная

$$e\rho^{(1)}(r,\ r) = e \sum_{p=1}^{k} |\psi_p(r)|^2, \tag{3,10}$$

может быть истолкована как пространственная плотность заряда электронов первого роя; аналогичное толкование допускает величина $e\rho^{(2)}(r,\ r)$. Выражения же (3,08) и (3,09) с разными аргументами r, r' не допускают классического толкования; их можно условно назвать «смешанной плотностью».

Переходим к толкованию выражения (3,07) для энергии системы электронов. Первая сумма представляет кинетическую и потенциальную энергию электронов первого роя в поле ядер; вторая сумма пред-

ставляет то же самое для второго роя. Те члены в двойном интеграле, которые содержат плотность $\rho^{(1)}$ от одинаковых аргументов, представляют электростатическую энергию электронов первого роя. Член же со смешанной плотностью не допускает классического толкования, и наличие его в выражении для энергии представляет специфически квантовый эффект (так называемая энергия квантового обмена). Второй двойной интеграл имеет тот же смысл для электронов второго роя. Наконец, последний двойной интеграл представляет взаимную электростатическую энергию обоих роев электронов.

Приведенное толкование наших формул хотя и не строго, но отличается наглядностью, и потому полезно для понимания их физического смысла. Строгое же толкование выражения (3,07) сводится к тому, что это выражение представляет результат подстановки в варьируемый интеграл волновой функции, обладающей надлежащей симметрией.

Система уравнений для искомых функций $\psi_p(r)$ получается путем вариации выражения (3,07) при добавочных условиях (3,04). Эта система имеет вид

$$2[H(r)+V(r)]\psi_p(r)-e^2\int\frac{[\rho^{(1)}(r',r)+\rho^{(2)}(r',r)]}{|\mathbf{r}-\mathbf{r}'|}\psi_p(r')\,d\tau'=$$
$$=\sum_{q=1}^{n-k}\lambda_{qp}\psi_q(r)\qquad(p=1,2,\ldots,k),\tag{3,11}$$

$$[H(r)+V(r)]\psi_p(r)-e^2\int\frac{\rho^{(2)}(r',r)}{|\mathbf{r}-\mathbf{r}'|}\psi_p(r')\,d\tau'=$$
$$=\sum_{q=1}^{n-k}\lambda_{qp}\psi_q(r)\qquad(p=k+1,\ldots,n-k).\tag{3,12}$$

В этих формулах величина $V(r)$ имеет значение

$$V(r)=e^2\int\frac{\rho^{(1)}(r',r')+\rho^{(2)}(r',r')}{|\mathbf{r}-\mathbf{r}'|}\,d\tau'\tag{3,13}$$

и может быть истолкована как умноженный на e потенциал всех электронов. Величины λ_{qp} суть лагранжевы множители, соответствующие условиям ортогональности (3,04), которые должны быть учтены при составлении вариации δW. Можно считать, что не-диагональные элементы матрицы λ_{qp} отличны от нуля только если один из значков больше или равен $k+1$, а другой меньше или равен k.

Заметим, что уравнение для функции $\psi_p(r)$ фактически не содержит этой функции в своих коэффициентах, так что если считать все остальные функции, кроме $\psi_p(r)$, известными, то уравнение для $\psi_p(r)$ будет линейным.

Это свойство уравнений можно формулировать следующим образом. Положим

$$\rho_p^{(1)}(r, r') = \sum_{q=1}^{k} (1 - \delta_{pq}) \overline{\psi}_q(r) \psi_q(r'), \qquad (3,14)$$

$$\rho_p^{(2)}(r, r') = \sum_{q=1}^{n-k} (1 - \delta_{pq}) \overline{\psi}_q(r) \psi_q(r') \qquad (3,15)$$

и обозначим через $V_p(r)$ выражение, аналогичное (3,13), но вычисленное при помощи сумм (3,14) и (3,15) вместо (3,08) и (3,09). Тогда уравнения (3,11) и (3,12) сохранят свой вид, если в каждом из них заменить V на V_p и $\rho^{(1)}$ и $\rho^{(2)}$ на $\rho_p^{(1)}$ и $\rho_p^{(2)}$.

Выпишем эти уравнения для того случая, когда на каждой орбите имеется по два электрона, т. е. когда $s=0$. Тогда n будет четным числом, $k=\frac{n}{2}$ и обе суммы (3,08) и (3,09) совпадут, так что мы можем отбросить верхний значок при ρ. Кроме того, в этом случае можно считать λ_{qp} диагональной матрицей и положить

$$\lambda_{qp} = 2E_p \delta_{qp}. \qquad (3,16)$$

В результате мы будем иметь

$$[H(r) + V_p(r)] \psi_p(r) - e^2 \int \frac{\rho_p(r', r)}{|\mathbf{r} - \mathbf{r}'|} \psi_p(r') d\tau' = E_p \psi_p(r), \qquad (3,17)$$

где

$$V_p(r) = 2e^2 \int \frac{\rho(r, r')}{|\mathbf{r} - \mathbf{r}'|} d\tau' - 2e^2 \int \frac{|\psi_p(r')|^2}{|\mathbf{r} - \mathbf{r}'|} d\tau'. \qquad (3,18)$$

Если в уравнении (3,17) отбросить интегральный член, то оно может быть истолковано как уравнение Шредингера для электрона в поле с потенциальной энергией

$$\Phi = U(r) + V_p(r), \qquad (3,19)$$

где $U(r)$ есть входящая в (3,06) потенциальная энергия внешнего поля, а $V_p(r)$ потенциальная энергия от всех прочих электронов, кроме

данного. Действительно, $V_p(r)$, вычисляемое по формуле (3,18), пропорционально потенциалу, соответствующему плотности $\rho - |\psi_p|^2$. Интегральный же член в уравнении (3,17) не может быть истолкован классически. Он получил название поправки на энергию квантового обмена.

Уравнения (3,17) без интегрального члена (вернее, несколько менее точные уравнения) были впервые предложены английским математиком Хартри, [1] который, однако, не дал им удовлетворительного обоснования, ибо не пользовался при их выводе вариационным началом и не рассматривал волновой функции системы электронов, а исходил из только что приведенных наглядных соображений. Они были названы им уравнениями согласованного поля (в том смысле, что потенциал V, входящий в уравнения для волновых функций, сам выражается через них).

Полные уравнения с интегральными членами, учитывающие свойства симметрии волновой функции системы для данного спина, были получены вскоре после того автором этой статьи [2], который вывел их из вариационного начала и тем самым обосновал также и уравнения Хартри. Они получили в литературе название уравнений согласованного поля с квантовым обменом. Формулированные в § 1 свойства симметрии координатной волновой функции также впервые найдены автором этой статьи [11].

Уравнения согласованного поля с квантовым обменом допускают и другую формулировку, указанную Дираком [2] и отличающуюся от изложенной тем, что спиновые переменные не исключаются с самого начала, а входят и в одноэлектронные волновые функции. Волновая функция системы (1,03) приближенно выражается через один определитель вида

$$\Psi = \begin{vmatrix} \psi_1(x_1) & \ldots & \psi_1(x_n) \\ \ldots & & \ldots \\ \psi_n(x_1) & \ldots & \psi_n(x_n) \end{vmatrix}, \tag{3,20}$$

[1] D. R. Hartree. The wave mechanics of an atom with a non-Coulomb central field. Proc. Cambr. Phil. Soc., 1928, vol. 24, p. 89, III.

[2] P. A. M. Dirac. Proc. Cambridge Phil. Soc. 1930, vol. 27, p. 240. Идея этого способа была независимо высказана также Слэтером в заметке: J. C. Slater. Note on Hartree's method. Phys. Rev. 1930, vol. 35, p. 210.

содержащий одноэлектронные функции (1,01), для которых и получаются уравнения согласованного поля, аналогичные нашим. Преимущество этого способа заключается в сравнительной простоте выкладок (так как приходится иметь дело с одним определителем, а не с произведением двух определителей); недостатком же является то, что уравнение (1,14) для оператора спинового момента количества движения выполняется не тождественно, а лишь при надлежащем выборе одноэлектронных функций. В случае сферической симметрии можно выразить входящие в определитель (3,20) функции $\psi_i(x)$ через радиальные функции R_{nl} и шаровые функции со спином так, чтобы уравнения для радиальных функций совпали с получаемыми по нашему первоначальному способу.

Уравнения согласованного поля для волновых функций со спином могут быть выведены также из теории вторичного квантования [12].

§ 4

УРАВНЕНИЕ ДЛЯ ВАЛЕНТНОГО ЭЛЕКТРОНА И ОПЕРАТОР КВАНТОВОГО ОБМЕНА

Рассмотрим систему, состоящую из нечетного числа $n=2k+1$ электронов и обладающую спином, равным $\frac{1}{2}$, например атом с одним валентным электроном. Полная волновая функция такой системы, имеющая вид произведения двух определителей, будет содержать k волновых функций $\psi_1, \psi_2, \ldots, \psi_k$, которые будут входить в оба определителя, и одну волновую функцию ψ, которая входит только в больший определитель. Мы можем сказать, что функции $\psi_1, \psi_2, \ldots, \psi_k$ описывают $2k$ внутренних электронов с компенсированным спином (по два на каждой „орбите“), а функция ψ описывает валентный электрон.

Пользуясь формулой (3,07), мы можем написать энергию такой системы в виде суммы

$$W=W_0+W', \qquad (4{,}01)$$

где

$$W_0 = 2\sum_{p=1}^{k}\int \overline{\psi_p}(r)\,H(r)\,\psi_p(r)\,d\tau + $$
$$+ e^2 \iint \frac{2\rho(r,r)\,\rho(r',r') - |\rho(r,r')|^2}{|\mathbf{r}-\mathbf{r}'|}\,d\tau\,d\tau' \tag{4,02}$$

есть энергия внутренних электронов и

$$W' = \int \overline{\psi}(r)\,H(r)\,\psi(r)\,d\tau + $$
$$+ e^2 \iint \frac{2\rho(r',r')\,|\psi(r)|^2 - \rho(r',r)\,\overline{\psi}(r)\,\psi(r')}{|\mathbf{r}-\mathbf{r}'|}\,d\tau\,d\tau' \tag{4,03}$$

есть энергия валентного электрона в поле внутренних электронов. Под $\rho(r, r')$ мы разумеем здесь смешанную плотность

$$\rho(r, r') = \sum_{p=1}^{k}\overline{\psi}_p(r)\,\psi_p(r'), \tag{4,04}$$

Если мы будем варьировать величину W по всем волновым функциям одновременно, мы вновь получим уравнения согласованного поля (3,11) и (3,12).

Но мы можем несколько видоизменить задачу и определить сперва волновые функции внутренних электронов из условия минимума величины W_0. После этого мы можем, считая $\psi_1, \psi_2, \ldots, \psi_k$ заданными, определить волновую функцию внешнего электрона $\psi(r)$ из условия минимума величины W или W' (последние две величины отличаются друг от друга на постоянную W_0).

Наша видоизмененная задача физически соответствует тому, что мы сперва определяем стационарное состояние системы, содержащей одним электроном меньше (атомный остов), а затем, пренебрегая поляризацией атомного остова валентным электроном, находим состояние этого последнего.

Для волновых функций внутренних электронов мы получаем систему уравнений (3,17), а для валентного электрона — линейное интегро-дифференциальное уравнение

$$[H(r) + V(r)]\,\psi(r) - e^2\int \frac{\rho(r',r)\,\psi(r')}{|\mathbf{r}-\mathbf{r}'|}\,d\tau' = E\psi(r), \tag{4,05}$$

где

$$V(r) = 2e^2 \int \frac{\rho(r', r')}{|\mathbf{r} - \mathbf{r}'|} d\tau'. \qquad (4,06)$$

Волновая функция валентного электрона все время предполагается ортогональной к волновым функциям внутренних электронов

$$\int \overline{\psi}_p(r)\psi(r)\,d\tau = 0 \qquad (p = 1, 2, \ldots, k). \qquad (4,07)$$

Согласно замечанию, сделанному выше по поводу уравнений (3,11) и (3,12), уравнение (3,17) не изменится, если в нем заменить величины ρ_p и V_p на ρ и V. Но при такой замене оно переходит в (4,05). Следовательно, уравнению для валентного электрона удовлетворяют также и все волновые функции внутренних электронов; все они являются собственными функциями одного и того же линейного интегро-дифференциального оператора, стоящего в левой части (4,05). Отсюда также следует, что условия ортогональности (4,07) выполняются сами собой (т. е. являются следствиями самого уравнения). Параметры же E_p, входящие в (3,17), являются собственными значениями того же оператора и могут быть истолкованы как уровни энергии внутренних электронов.

Введем линейный интегральный оператор A, определив его по равенствам

$$A\psi(r) = e^2 \int \frac{\rho(r', r)}{|\mathbf{r} - \mathbf{r}'|} \psi(r')\,d\tau'. \qquad (4,08)$$

Припоминая значение (3,06) оператора $H(r)$, мы можем написать наше основное уравнение (4,05) в виде

$$-\frac{h^2}{2m}\Delta\psi + [U(r) + V(r)]\psi - A\psi = E\psi. \qquad (4,09)$$

Оператор $(-A)$ входит слагаемым в выражение для энергии электрона. Его можно поэтому толковать как особый вид энергии, который принято называть энергией квантового обмена. Из вывода уравнения (4,09) следует, что наличие члена $A\psi$ связано с учетом свойств симметрии волновой функции, а эти свойства, как и принцип Паули, связаны с неотличимостью электронов друг от друга и с невозможностью просле-

дить за отдельным электроном, когда он вступает в тесное взаимодействие с другими электронами (невозможность наклеить на электрон ярлычок, чтобы опознать его после взаимодействия). Эта невозможность имеет место только в квантовой, но не в классической механике (где понятие траектории предполагается неограниченно применимым). Поэтому трудно придумать для оператора A наглядное толкование и название. Принятое название «квантовый обмен» связано с представлением, что электроны при тесном взаимодействии как бы обмениваются местами.

§ 5

ПРИМЕНЕНИЕ МЕТОДА СОГЛАСОВАННОГО ПОЛЯ К ТЕОРИИ СТРОЕНИЯ АТОМА

Наиболее простой многоэлектронной системой является атом. Применение метода согласованного поля к расчету различных атомов облегчается тем, что еще до возникновения современной формы квантовой механики (теории Шредингера) Бором была дана схема строения электронных оболочек всех атомов, входящих в периодическую систему элементов Менделеева. На основании этой схемы оказывается возможным приписать каждому электрону в атоме определенные квантовые числа, аналогичные тем, какими характеризуются в задаче одного тела состояния отдельного электрона в поле со сферической симметрией. Схема Бора была разработана им на экспериментальной основе, а именно на основании анализа спектров и химических свойств атомов. Свое теоретическое обоснование схема Бора получила с появлением квантовой механики Шредингера и с развитием приближенных методов квантовой механики, в особенности метода согласованного поля.

На языке квантовой механики возможность приписать каждому электрону в атоме определенные квантовые числа означает возможность приписать ему определенную волновую функцию. Это есть как раз то предположение, которое лежит в основе метода согласованного поля, где волновая функция всей системы выражается через волновые функции отдельных электронов. Знание же квантовых чисел данного электрона позволяет определить общий характер волновой функции.

По схеме Бора электроны в атоме разбиваются на группы эквивалентных электронов. Каждая такая группа характеризуется двумя квантовыми числами: главным квантовым числом n и азимутальным квантовым числом l, причем $n = 1, 2, 3, \ldots$ и $l = 0, 1, \ldots, n-1$. Внутри группы электроны могут быть охарактеризованы двумя другими квантовыми числами: магнитным квантовым числом m и спиновым квантовым числом m_s, причем m принимает значения $m = -l, -l+1, \ldots, l-1, l$, а m_s два значения: $m_s = \pm\frac{1}{2}$. В изложенной выше теории (§ 3) мы учитываем два значения спинового квантового числа тем, что подразделяем электроны на два роя. Поэтому для характеристики волновой функции электрона с данными n и l нам достаточно указать значение магнитного квантового числа m.

Если обозначить через r, ϑ, φ сферические координаты с началом в ядре, то волновая функция электрона с квантовыми числами n, l, m будет иметь вид

$$\psi_{nlm} = \frac{1}{\sqrt{4\pi}} R_{nl}(r) Y_{lm}(\vartheta, \varphi), \tag{5,01}$$

где Y_{lm} — шаровая функция, нормированная так, что

$$\int |Y_{lm}(\vartheta, \varphi)|^2 \sin\vartheta \, d\vartheta \, d\varphi = 4\pi. \tag{5,02}$$

Необходимо отметить, что квантовые числа, которые служат для подразделения электронов внутри группы, имеют условный характер вследствие произвола в выборе полярной оси. Этот произвол не сказывается, однако, в том случае, когда группа эквивалентных электронов заполнена целиком, т. е. когда в ней представлены электроны со всеми возможными значениями m. В таком случае говорят, что мы имеем замкнутую электронную оболочку.

Напишем выражение для смешанной плотности электронов одного роя, входящих в замкнутую электронную оболочку. По определению (4,04), мы будем иметь

$$\rho_{nl}(\mathbf{r}, \mathbf{r}') = \sum_{m=-l}^{+l} \bar{\psi}_{nlm}(r, \vartheta, \varphi) \psi_{nlm}(r', \vartheta', \varphi'). \tag{5,03}$$

Подставляя сюда значение ψ_{nlm} из (5,01) и пользуясь теоремой сложения для шаровых функций, мы будем иметь

$$\rho_{nl}(\mathbf{r}, \mathbf{r}') = \frac{2l+1}{4\pi} R_{nl}(r) R_{nl}(r') P_l(\cos\gamma), \tag{5,04}$$

где

$$\cos\gamma = \cos\vartheta\cos\vartheta' + \sin\vartheta\sin\vartheta'\cos(\varphi - \varphi'), \tag{5,05}$$

а P_l есть полином Лежандра. Так как γ есть угол между направлениями (ϑ, φ) и (ϑ', φ'), то он не зависит от выбора полярной оси. Формула (5,04) позволяет заключить, что замкнутая электронная оболочка обладает сферической симметрией.

Число электронов одного роя, входящих в замкнутую оболочку, равно $2l+1$, и, следовательно, полное число электронов такой оболочки равно $4l+2$.

Согласно схеме Бора, большинство электронов атома входит в состав замкнутых оболочек. Эти электроны можно назвать внутренними. Остальные электроны, которые можно назвать внешними, располагаются в незаполненных оболочках. За исключением редких земель и немногих других элементов, незаполненной остается только одна оболочка, а в этих исключительных случаях — две оболочки.

Наиболее простым является тот случай, когда вне заполненных оболочек остается только один (валентный) электрон. В качестве примера такого атома мы рассмотрим атом натрия.

При решении конкретных задач, относящихся к атому, удобно пользоваться введенными Хартри атомными единицами, в которых заряд, масса электрона и деленная на 2π постоянная Планка полагаются равными единице. Атомной единицей длины будет тогда 0,53 ангстрема, атомной единицей скорости $\frac{1}{137}$ скорости света, а атомной единицей энергии — удвоенный потенциал ионизации водорода, т. е. 27,09 электрон-вольта.

В атомных единицах оператор $H(\mathbf{r})$ для натрия будет равен

$$H(\mathbf{r}) = -\frac{1}{2}\Delta - \frac{11}{r}, \tag{5,06}$$

так как заряд ядра натрия равен 11 единицам. Внутренние электроны атома натрия располагаются в трех замкнутых оболочках. Соответствующие волновые функции мы можем положить равными

$$\psi_1 = \psi_{10,0}, \qquad \psi_2 = \psi_{20,0}, \qquad \psi_3 = \psi_{21,-1}, \\ \psi_4 = \psi_{21,0}, \qquad \psi_5 = \psi_{21,1}, \tag{5,07}$$

где ψ_{nlm} имеют значение (5,01). Смешанная плотность всех трех оболочек будет иметь вид

$$\rho(\mathbf{r}, \mathbf{r}') = \frac{1}{4\pi}[R_{10}(r) R_{10}(r') + R_{20}(r) R_{20}(r') + 3 R_2 (r) R_{21}(r') \cos\gamma]. \tag{5,08}$$

Подставляя эти выражения в (4,02) для W_0, мы получим энергию ионизованного атома натрия. В полученном выражении для W_0 мы можем произвести интегрирование по углам, и тогда оно будет содержать только радиальные функции R_{10}, R_{20} и R_{21}. Дальнейшие вычисления можно вести двумя способами: численно или аналитически. Для численного решения необходимо составить вариационные уравнения для радиальных функций; они будут иметь вид аналогичный (3,17). Методика численного решения вариационных уравнений нами подробно разработана [3], она основана на способе последовательных приближений и позволяет составлять таблицы радиальных функций с желаемой степенью точности. Для успешного ее применения необходимо иметь хорошее начальное приближение, которое лучше всего получать аналитическим путем. Для этого нужно искать R_{10}, R_{20}, R_{21} в виде аналитических выражений,[1] содержащих небольшое число параметров. Мы можем взять, например

$$R_{10} = ae^{-\alpha r}, \quad R_{20} = b\left[1 - \frac{1}{3}(\alpha + \beta) r\right] e^{-\beta r}, \quad R_{21} = cre^{-\gamma r}, \tag{5,09}$$

где a, b, c определяются из условия нормировки

$$\int_0^\infty [R_{nl}(r)]^2 r^2 dr = 1 \tag{5,10}$$

[1] Аналитические выражения для волновых функций были построены нами также для атомов типа бериллия [5], содержащих два валентных электрона, причем для уровней энергии и потенциалов ионизации получились значения, близкие к экспериментальным.

и выражаются через α, β, γ. Подставляя (5,09) в выражение для W_0, мы получим W_0 как дробную рациональную функцию от α, β, γ. Приравнивая нулю производные от W_0 по параметрам α, β, γ, мы приходим к уравнениям, из которых эти параметры и определяются. В нашем случае мы будем иметь

$$\alpha = 10{,}68, \qquad \beta = 4{,}22, \qquad \gamma = 3{,}49, \tag{5,11}$$

и наши аналитические радиальные функции будут иметь вид

$$R_{10}(r) = 69{,}804\, e^{-10.68\, r}, \qquad R_{20}(r) = 13{,}602\,(1-4{,}967\, r)\, e^{-4.22\, r};$$
$$R_{21}(r) = 26{,}276\, r\, e^{-3.49\, r}. \tag{5,12}$$

Беря эти функции в качестве исходного приближения, можно получить более точные их значения из вариационных уравнений. Функциям (5,12) соответствует значение энергии ионизованного атома натрия, равное $W_0 = -160{,}9$ атомных единиц, тогда как более точным функциям, полученным в результате численного решения уравнений, соответствует число $W_0 = -161{,}8$.

Найдя радиальные функции R_{10}, R_{20} и R_{21}, мы тем самым полностью определили в данном приближении структуру внутренних электронных оболочек атома натрия. После этого мы можем составить интегро-дифференциальное уравнение (4,05) для волновой функции валентного электрона. Эта функция будет иметь вид (5,01), где $R_{nl}(r)$ получается в результате численного интегрирования соответствующего уравнения, которого мы не будем здесь выписывать.

Решение уравнения (4,05) дает прежде всего оптические термы атома натрия. Кроме того, оно дает, правда довольно грубо, и значения рентгеновских термов. Чтобы судить о получаемой степени точности, мы приведем следующую таблицу:

	E_{10}	E_{20}	E_{21}	E_{30}	E_{31}
Теория	−40,6	−3,00	−1,83	−0,1860	−0,1094
Эксперимент	−39,4	−2,16	−1,04	−0,18885	−0,1115

Все числа выражены в атомных единицах. Из таблицы видно, что оптические термы (E_{30} и E_{31}) получаются с довольно большой степенью точности. Так, погрешность в терме E_{31} составляет только 1,9%. Для сравнения укажем, что если отбросить в уравнении (4,05) интегральный член, то значение терма E_{31} получается равным $-0{,}08895$, что соответствует погрешности в 22,4%.

Кроме оптических термов, наше интегро-дифференциальное уравнение дает и волновые функции валентного электрона для разных уровней, что позволяет вычислить и вероятности переходов. Подобные вычисления были произведены нами для атомов натрия и лития [2, 3, 6], причем получилось качественное согласие с опытом. В частности, в случае лития получился характерный для этого элемента немонотонный ход зависимости вероятности перехода $E_{n1} \to E_{20}$ от квантового числа n.

Изложенные выше результаты показывают, что теория валентного электрона в атоме действительно может быть с большой точностью формулирована как задача одного тела в заданном поле, если только учитывать силы квантового обмена при помощи нашего интегро-дифференциального уравнения (4,05) [4].

Одно из применений нашего уравнения относится к вопросу о так называемом правиле сумм [7]. Правило это относится к так называемым «силам осцилляторов», т. е. величинам, пропорциональным квадрату матричного элемента для координаты, соответствующего данному переходу, и разности уровней энергии для этого перехода. Согласно обычной формулировке этого правила, сумма сил осцилляторов, соответствующих всем оптическим переходам данной серии, должна равняться единице. Это правило строго соблюдается для атомов с одним электроном (атомы, подобные атому водорода). Что же касается других атомов с одним валентным электроном, то опыт показывает, что в ряде случаев (литий, таллий и др.) уже первые «осцилляторы» дают в сумме число, большее единицы. Наша теория дает этому обстоятельству объяснение, основанное на том, что уравнение (4,05) имеет замкнутую систему собственных функций, причем некоторые из них соответствуют рентгеновским уровням. Поэтому при составлении суммы сил осцилляторов необходимо учитывать также и те фиктивные осцилляторы, которые относятся к переходам на занятые рентгеновские (более низкие) уровни.

Так как для них входящая в выражение для «силы» разность уровней отрицательна, то и «сила» их будет отрицательной. Поэтому в выражение для полной суммы будет входить и несколько (конечное число) отрицательных членов. Отсюда ясно, что если вся сумма равна единице, то сумма положительных членов (которые соответствуют наблюдаемым оптическим переходам) будет больше единицы. Кроме того, нужно иметь в виду, что и вся сумма будет равна единице, только если в ее выражении пренебрегать некоторыми малыми поправками, связанными с квантовым обменом.

Основные применения наших уравнений к атомам с одним валентным электроном относятся к расчету уровней и интенсивностей. Кроме того, были сделаны попытки учесть некоторые релятивистские эффекты, а именно расстояние между термами дублета (дублетное расщепление) [8].

Вследствие трудностей, связанных с определением волновой функции валентного электрона на малых расстояниях от ядра, попытки эти не дали хорошего численного совпадения с опытом, а дали для дублетного расщепления термов натрия и лития только правильный порядок величины. (Заметим, что прежние попытки, основанные на более грубом приближении для волновой функции, не давали даже правильного порядка). При этом выяснилась важная роль, которую играют в формуле для дублетного расщепления члены, дающие поправки на квантовый обмен. Не исключено, что эти члены позволят объяснить наблюдаемое для некоторых атомов отрицательное значение величины дублетного расщепления. Представляется, однако, весьма вероятным, что для построения более точной теории дублетного расщепления рассмотрение валентного электрона в фиксированном поле внутренних электронов окажется недостаточным.

§ 6

ТЕОРИЯ ВОЛНОВЫХ ФУНКЦИЙ ВОДОРОДНОГО ТИПА

В настоящем параграфе мы рассмотрим основные идеи предложенной нами [9] теории волновых функций водородного типа, позволяющей изучать свойства совокупности этих функций, относящихся к данному значению главного квантового числа n.

По отношению к содержанию предыдущих параграфов теория таких функций стоит несколько особняком, но одно из ее приложений относится к строению атомных оболочек, относящихся к определенному значению главного квантового числа n.

В предыдущем параграфе мы вместе с Бором характеризовали каждую атомную оболочку двумя квантовыми числами n и l. При данном n число l может принимать значения $l=0, 1, 2, \ldots, n-1$, всего n значений. Совокупность всех n электронных оболочек, принадлежащих данному n, образует одну «большую оболочку». Такая «большая» оболочка обладает в атоме особой устойчивостью. В одновалентных атомах лития, натрия и меди полностью заполнены соответственно одна, две и три «большие» оболочки.

Большую оболочку удобно описывать при помощи волновых функций водородного типа, соответствующих некоторому эффективному заряду ядра, значение которого можно определить из вариационного начала.

Обозначим через Z истинный заряд ядра и через Z_n эффективный заряд для большой оболочки с квантовым числом n. Вместо эффективного заряда Z_n удобно рассматривать величину $p_n = \frac{Z_n}{n}$, представляющую среднее квадратичное количество движения электрона в большой оболочке номер n (в атомных единицах). Для атомов Na^+, Al^{+++}, Cu^+, Zn^{++} (где число крестиков обозначает степень ионизации атома) получаются следующие значения p_n:

Na^+	$(Z=11)$	$p_1=10,63$	$p_2=3,45$	
Al^{+++}	$(Z=13)$	$p_1=12,62$	$p_2=4,45$	
Cu^+	$(Z=29)$	$p_1=28,59$	$p_2=10,64$	$p_3=5,47$
Zn^{++}	$(Z=30)$	$p_1=29,59$	$p_2=11,09$	$p_3=5,84$

Описание атомных оболочек при помощи аналитических функций водородного типа представляет большие преимущества, так как позволяет находить простые аналитические выражения для различных функций, характеризующих свойства атома. Например, функция рас-

пределения количества движения электронов внутри n-ой оболочки, нормированная согласно условию

$$4\pi\int_0^\infty \rho_n(p)\,p^2\,dp = n^2, \tag{6,01}$$

оказывается равной

$$\rho_n(p) = \frac{8p_n^5 n^2}{\pi^2(p_n^2+p^2)^4}. \tag{6,02}$$

Далее так называемый фактор атомной формы для n-ой оболочки, определяемый равенством

$$F_n = \int e^{i\mathbf{k}.\mathbf{r}} \sum_{l=0}^{n-1} \rho_{nl}(\mathbf{r},\ \mathbf{r})\,d\tau, \tag{6,03}$$

где ρ_{nl} имеет значение (5,04), получается равным

$$F_n = F_n(x) = \frac{1}{4n^2}\,T'_n(x)\,(1+x)^2\,[P'_n(x)+P'_{n-1}(x)], \tag{6,04}$$

где положено

$$x = \frac{4p_n^2-k^2}{4p_n^2+k^2} \tag{6,05}$$

и через $P'_n(x)$ и $T'_n(x)$ обозначены производные от полинома Лежандра $P_n(x)$ и полинома Чебышева

$$T_n(x) = \cos(n \arccos x). \tag{6,06}$$

В частности, при $n=1, 2, 3$ будет

$$F_1(x) = \frac{1}{4}(1+x)^2, \qquad F_2(x) = \frac{1}{4}(1+x)^2 x(3x+1),$$

$$F_3(x) = \frac{1}{8}(1+x)^2(5x^2+2x-1)(4x^2-1). \tag{6,07}$$

Помимо применений к теории строения электронных оболочек, теория водородных функций, принадлежащих данному квантовому числу n, может иметь применение в теории явления Комптона от связанных электронов и в других аналогичных задачах, в которых

приходится иметь дело с функциями, принадлежащими сплошному спектру. Имея дело с аналитическими функциями от квантового числа n, можно формально переходить от точечного спектра к сплошному, придавая величине n чисто мнимые значения. Поэтому мы ограничимся, в дальнейшем случаем точечного спектра.

Напишем уравнение Шредингера для водородоподобного атома в пространстве импульсов. Так как оператор умножения на кулоновскую потенциальную энергию $-\frac{Ze^2}{r}$ является в пространстве импульсов интегральным оператором, то уравнение Шредингера будет интегральным уравнением вида

$$\frac{1}{2m}\,p^2\psi(\mathbf{p}) - \frac{Ze^2}{2\pi^2 h}\int\frac{\psi(\mathbf{p}')\,(d\mathbf{p}')}{|\,\mathbf{p}-\mathbf{p}'\,|^2} = E\psi(\mathbf{p}), \tag{6,08}$$

где через $(d\mathbf{p}') = dp'_x\,dp'_y\,dp'_z$ обозначен элемент объема в пространстве импульсов. Так как мы рассматриваем точечный спектр, для которого энергия E отрицательна, то можно ввести средний квадратичный импульс

$$p_0 = \sqrt{-2mE}. \tag{6,09}$$

Деленные на p_0 составляющие вектора количества движения $\mathbf{p}$ мы будем рассматривать как прямоугольные координаты на гиперплоскости, представляющей стереографическую проекцию шара радиуса единицы в четырехмерном евклидовом пространстве. Прямоугольные координаты некоторой точки на шаре будут

$$\begin{aligned}
\xi &= \frac{2p_0 p_x}{p_0^2 + p^2} = \sin\alpha\sin\vartheta\cos\varphi,\\
\eta &= \frac{2p_0 p_y}{p_0^2 + p^2} = \sin\alpha\sin\vartheta\sin\varphi,\\
\zeta &= \frac{2p_0 p_z}{p_0^2 + p^2} = \sin\alpha\cos\vartheta,\\
\chi &= \frac{p_0^2 - p^2}{p_0^2 + p^2} = \cos\alpha,
\end{aligned} \tag{6,10}$$

причем

$$\xi^2+\eta^2+\zeta^2+\chi^2=1. \qquad (6,11)$$

Углы α, ϑ и φ представляют сферические координаты на гиперсфере. Вместе с тем углы ϑ и φ являются обыкновенными сферическими углами, характеризующими направление количества движения. Элемент поверхности на гиперсфере равен

$$d\Omega=\sin^2\alpha\, d\alpha \sin\vartheta d\vartheta\, d\varphi, \qquad (6,12)$$

а полная поверхность гиперсферы равна $2\pi^2$. Введем вместо $\psi(\mathbf{p})$ функцию

$$\Psi(\alpha,\ \vartheta,\ \varphi)=\frac{\pi}{\sqrt{8}}\, p_0^{-\frac{5}{2}}(p_0^2+p^2)^2\,\psi(\mathbf{p}), \qquad (6,13)$$

для которой условие нормировки имеет вид

$$\frac{1}{2\pi^2}\int|\Psi(\alpha,\ \vartheta,\ \varphi)|^2 d\Omega=\int\frac{p^2+p_0^2}{2p_0^2}|\psi(\mathbf{p})|^2(d\mathbf{p})=\int|\psi(\mathbf{p})|^2(d\mathbf{p})=1. \qquad (6,14)$$

Если мы положим для краткости

$$\lambda=\frac{Zme^2}{hp_0}=\frac{Zme^2}{h\sqrt{-2mE}} \qquad (6,15)$$

и перейдем к новым переменным, то уравнение Шредингера (6,08) примет вид

$$\Psi(\alpha,\ \vartheta,\ \varphi)=\frac{\lambda}{2\pi^2}\int\frac{\Psi(\alpha',\ \vartheta',\ \varphi')}{4\sin^2\frac{\omega}{2}}\,d\Omega'. \qquad (6,16)$$

Здесь $2\sin\frac{\omega}{2}$ есть длина хорды, а ω — длина дуги большого круга, соединяющей точки α, ϑ, φ и α', ϑ', φ' на четырехмерном шаре, так что

$$4\sin^2\frac{\omega}{2}=(\xi-\xi')^2+(\eta-\eta')^2+(\zeta-\zeta')^2+(\chi-\chi')^2. \qquad (6,17)$$

Уравнение (6,17) представляет собою не что иное, как интегральное уравнение для шаровых функций четырехмерного шара. Собственными значениями будут целые числа $\lambda=n$ ($n=1, 2, \ldots$), а собственными функциями будут однородные гармонические полиномы степени $n-1$ от ξ, η, ζ, χ, т. е. функции вида

$$\Psi=u(\xi, \eta, \zeta, \chi), \tag{6,18}$$

где $u(x_1, x_2, x_3, x_4)$ есть решение уравнения Лапласа в четырехмерном пространстве

$$\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\frac{\partial^2 u}{\partial x_3^2}+\frac{\partial^2 u}{\partial x_4^2}=0. \tag{6,19}$$

Как видно из (6,15), целое число n есть главное квантовое число.

Таким образом, теория атома водорода оказывается связанной с четырехмерной теорией потенциала. Связь эта позволяет легко вывести все свойства водородных функций и, в частности, установить для них теорему сложения, справедливую не только для целых действительных значений n (точечный спектр), но и для комплексных n (сплошной спектр).

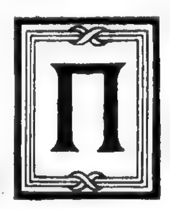
Наиболее существенным следствием наличия такой связи является установление группы преобразований, допускаемых уравнением Шредингера для атома водорода. Очевидно, что уравнение (6,17) сохранит свой вид, если произвести над переменными ξ, η, ζ, χ ортогональную подстановку, т. е. если подвергнуть гиперсферу произвольному четырехмерному вращению. Отсюда следует, что и исходное уравнение Шредингера обладает не только обычной сферической симметрией, но и более высокой степенью симметрии, соответствующей четырехмерным вращениям. Этим объясняется тот давно известный факт, что уровни энергии для водорода зависят только от главного квантового числа n. Использование такой более широкой группы преобразований уравнения Шредингера и позволяет получить те результаты, о которых мы говорили в начале этого параграфа. На подробной формулировке этих результатов, а также на их выводе мы здесь останавливаться не можем.

ПЕРЕЧЕНЬ РАБОТ АВТОРА, УПОМИНАЕМЫХ ИЛИ ИСПОЛЬЗОВАННЫХ В НАСТОЯЩЕЙ СТАТЬЕ

1. *В. Фок.* Приближенный способ решения квантовой задачи многих тел, Тр. Гос. Оптич. ин-та, т. 5, вып. 51. Л. 1931.
2. *В. Фок.* Применение обобщенного способа Хартри к атому натрия. Тр. Гос. Оптич. ин-та, т. 5, вып. 51, Л., 1931.
3. *В. Фок* и *М. Петрашень.* О численном решении уравнений согласованного поля. ЖЭТФ, 1934, т. 4, стр. 295.
4. *В. Фок.* Об энергии квантового обмена. ЖЭТФ, 1934, т. 4, стр. 5.
5. *В. Фок* и *М. Петрашень.* Аналитические волновые функции атомов типа бериллия. ЖЭТФ, 1936, т. 6, стр. 1.
6. *V. Fock* a. *M. Petrashen.* Self-consistent field with exchange for lithium. Sow. Phys. 1935, vol. 8, p. 547.
7. *V. Fock.* Über die Anwendbarkeit des quantenmechanischen Summensatzes, Zs. f. Phys., Bd. 89, 1934, S. 744.
8. *В. Фок.* О дублетном расщеплении термов щелочных металлов. ЖЭТФ, 1935, т. 5, стр. 891.
9. *В. Фок.* Атом водорода и не-евклидова геометрия. Изв. АН СССР, отд. мат. и естеств. наук, 1935, стр. 169.
10. *В. Фок.* Новый вывод векторной модели. ЖЭТФ, 1940, т. 10, стр. 383.
11. *В. Фок.* О волновых функциях многоэлектронных систем. ЖЭТФ, 1940, т. 10, стр. 961.
12. *V. Fock.* Konfigurationsraum und zweite Quantelung, Zs. f. Phys., Bd. 75, 1932, S. 622.

Академик
И. В. Курчатов
и доктор физико-математических наук
Л. И. Русинов

ИЗОМЕРИЯ АТОМНЫХ ЯДЕР



Периодическим законом Менделеева впервые в наиболее определенной форме была установлена закономерная связь между химическими элементами, встречающимися в природе. В последующие десятилетия многочисленные исследования химических свойств элементов, огромный материал по изучению оптических и рентгеновских атомных спектров и, наконец, теоретические обобщения новой волновой механики позволили создать полную теорию электронной оболочки атомов и выяснить тонкости ее структуры.

Было показано, что индивидуальность химических элементов обусловлена прежде всего свойствами их атомных ядер. Положительный заряд ядра связывает равные количеству внутриядерных протонов орбитальные электроны и тем определяет основные характеристики атомов, а также порядковый номер элемента в периодической системе.

Опыты по изучению превращений естественно радиоактивных элементов, исследования в области масс-спектрографии и бурное развитие ядерной химии установили, что каждый элемент периодической системы может иметь большое число различных устойчивых и распадающихся изотопов.

В настоящее время таблица химических элементов, в которую входят все известные стабильные и радиоактивные изотопы, весьма обширна.

В середине 1944 г., по данным Сиборга, в нее входило свыше 700 различных изотопов. Их число сейчас превышает 800, и оно непрерывно расширяется за счет новых ядерных реакций. Искусственным образом удалось также увеличить и число элементов, создав элементы с атомными номерами до 96 включительно.

При детальном рассмотрении таблицы изотопов обращает на себя внимание то обстоятельство, что в довольно значительном числе встречаются такие атомные ядра, которые при одинаковых атомных и массовых числах обладают различными физическими свойствами, например имеют различные периоды распада, в частности, одни радиоактивны, а другие устойчивы. Такого рода изотопы с совпадающим числом протонов и нейтронов в ядре, но с отличающимися физическими характеристиками названы ядерными изомерами.

Вопросу ядерной изомерии и посвящена настоящая статья.

О т к р ы т и е я д е р н о й и з о м е р и и б р о м а. В 1934 г., всего лишь через два года после открытия нейтрона, Ферми удалось путем воздействия нейтронами на атомные ядра вызвать искусственную радиоактивность почти у всех элементов периодической системы. В частности, им была обнаружена искусственная радиоактивность ядер атомов брома, которая уменьшалась с двумя периодами полураспада: 18 мин. и 4.2 часа. Это находилось в полном соответствии с известными масс-спектрографическими данными о наличии у брома двух стабильных изотопов с массовыми числами 79 и 81. Предполагалось, что в результате захвата нейтронов образуются радиоактивные изотопы брома Br^{80} и Br^{82}, которые путем бета-распада с указанными выше периодами превращаются в устойчивые ядра криптона.

В начале 1935 г. авторами настоящей статьи в совместной работе с Б. В. Курчатовым и Л. В. Мысовским был открыт еще один, третий, период бета-радиоактивного распада ядер брома, облученного нейтронами.

При изучении на счетчике Гейгер-Мюллера гамма-излучения, сопровождающего бета-превращения брома, была замечена остаточная активность, период полураспада которой оказался равным приблизительно 36 час. Естественно возник вопрос, распаду какого радиоактивного ядра следует приписать этот новый период бета-излучения. Можно было предположить, что у брома имеется еще один, третий, стабильный изотоп,

не открытый еще масс-спектрографическим путем, который и обусловливает появление новой константы распада.

В связи с нашей работой Блевет в Америке предпринял экспериментальное исследование с целью обнаружить у брома еще один устойчивый изотоп.

На основании его опытов, выполненных с помощью масс-спектрографа, можно было заключить, что в пределах массовых чисел, от 73 до 87 включительно, у брома нет стабильных изотопов с распространенностью большей, чем 1/3000 общего числа атомов брома. Согласно последним данным Вильямса и Юстера (1946 г.), содержание стабильных изотопов брома с массовым числом, отличным от 79 и 81, не превышает 1/10 000.

Следовательно, не представлялось возможным объяснить появление третьего радиоактивного периода распада брома за счет наличия еще одного устойчивого изотопа.

Не исключалось, что вновь открытый период распада следует приписать радиоактивным ядрам, которые получаются в результате реакций с вылетом из ядра другой частицы.

Можно было думать, что имеет место, например, ядерная реакция с вылетом протона (n, p). Тогда радиоактивными должны были бы быть соседние с бромом по таблице Менделеева атомы селена. Тщательные химические опыты, проведенные Б. В. Курчатовым, установили, что бета-активности не было в тех препаратах, которые извлекались из бромистого этила вместе с атомами селена.

Отпало и предположение о том, что радиоактивность брома обусловлена реакцией с вылетом альфа-частицы (n, α). В этом случае радиоактивность извлекалась бы с атомами мышьяка. Однако опыты Б. В. Курчатова показали, что радиоактивности нет в химически выделенном мышьяке и она всегда обнаруживается в препарате, который химически выделяется с бромом. Оставалась возможность приписать третий период радиоактивности изотопу брома, который мог бы быть получен за счет реакции с вылетом нейтрона $(n, 2n)$. Л. И. Русиновым были предприняты опыты с целью проверки этого предположения.

Реакции $(n, 2n)$ имеют энергетический порог. Исследования Русинова показали, однако, что ядерная реакция, вызывающая активность брома с полупериодом в 36 час., не имеет такого порога. Этим было доказано,

что бета-активность брома с новым периодом, как и первые две, определяется процессами захвата нейтронов (реакция n, γ).

Весь вопрос в целом о причине возникновения у брома трех периодов радиоактивного распада при наличии двух стабильных изотопов оставался открытым, и необходимо было привлечь новые гипотезы для его объяснения.

В 1936 г. И. В. Курчатов в своем докладе на мартовской сессии Академии Наук СССР произвел детальное рассмотрение вопроса об аномалии в радиоактивности брома и пришел к заключению: «Таким образом, если не делать новых допущений, остается предположить, что в противоречии с данными масс-спектрографического анализа у брома имеется еще третий, мало распространенный устойчивый изотоп, путем присоединения к которому нейтрона и возникает третье радиоактивное ядро. Вместе с этим, крайне мало вероятным, предположением необходимо сделать и другое допущение, также весьма мало вероятное. Число ядер трех радиоактивных бромов — одного и того же порядка, и, следовательно, третий устойчивый изотоп, очень мало распространенный, должен обладать очень большим сечением для захвата нейтрона. Возможен и другой путь объяснения происхождения третьего радиоактивного брома, путь, который, правда, приводит к существенно новому допущению о существовании ядер изомеров. Говоря об изомерии ядра, не следует трактовать ее в том узком смысле, который был указан Гамовым. Ядра-изомеры в нашем случае — это ядра с одним и тем же зарядом, одним и тем же массовым числом (не различимые поэтому масс-спектрографически), но разной структуры».

Здесь с определенностью было высказано предположение, что в случае брома имеет место новое физическое явление — изомерия искусственно радиоактивных атомных ядер.

Вопрос об аналогичных аномалиях в области радиоактивности привлекает внимание и других исследователей. Известный физик Л. Майтнер в своем докладе на происходившем в июне 1936 г. в Цюрихе физическом съезде заявила:

«В настоящее время трудно поверить в существование «изомерных» атомных ядер, т. е. таких ядер, которые при равном атомном весе и равном атомном номере обладают различными радиоактивными свойствами. В связи с этим, может быть, будет интересно отметить, что и при других

элементах (кроме искусственно радиоактивного урана.— *К.* и *Р.*) мы до сих пор не знаем отчетливых переходов при воздействии нейтронов.

Например, у брома доказаны три различных радиоактивных процесса, которые усиливаются парафином, хотя существуют только два изотопа брома; аналогично складываются обстоятельства у индия и, может быть, у родия. Далее, в ряде урана UZ известен как продукт превращения, который следует рассматривать как изобарный и изотопный с UX_{II}, несмотря на то, что он обладает совершенно другой скоростью распада.

Предположение о существовании „изомерных" ядер дало бы возможность объяснить искусственные превращения урана. Оно ведет также к известным следствиям, которые можно доказать экспериментально. Мы надеемся после проведения этих экспериментов узнать, следует ли серьезно заниматься вопросом об „изомерных" ядрах».

В этом докладе, правда в более осторожной форме, также высказывается предположение о возможности существования нового класса ядер — ядерных изомеров. Следует заметить, что, хотя в течение ряда лет был известен единственный случай так называемой изотопии высшего порядка в ряду естественно радиоактивного урана (UZ и UX_{II}), открытие ядерной изомерии у искусственно радиоактивных элементов дало возможность начать поиски и значительно увеличить число известных случаев изомерных ядер, расширить объекты и пути экспериментальных исследований и направить их на выяснение физической сущности ядерной изомерии.

Первые теоретические представления о природе ядерной изомерии. Первая попытка теоретического толкования явления ядерной изомерии относится к 1934 г. Для объяснения аномалии в естественно радиоактивном ряду урана была выдвинута идея о возможности существования в составе некоторых атомных ядер гипотетических частиц — антипротонов, имеющих массу протона, но обладающих отрицательным зарядом. Однако эта теория не получила развития, хотя в ней и заключался ряд общих положений, которые были приняты в дальнейших теориях изомерии атомных ядер.

В конце 1936 г. Бор и Вайцзекер для объяснения известных к тому времени случаев изомерного брома и других элементов выдвинули новую теорию ядерной изомерии. Согласно этой теории, два ядерных изомера

обладают одинаковым составом протонов и нейтронов, но отличаются между собой тем, что находятся в двух различных энергетических состояниях. Обычно возбужденное ядро, обладающее бо́льшим запасом энергии, за весьма короткое время переходит путем электромагнитного гамма-излучения в нормальное ядро с меньшей энергией. Однако в случае изомерных ядер вследствие квантово-механических законов запрета такое излучение делается мало вероятным, и по этой теории ядра-изомеры находятся: одно в метастабильном, другое в основном состояниях. Причина метастабильности заключается в большой разности угловых моментов изомерных ядер.

Данные теории показывают, что вероятность высвечивания возбужденного изомера может быть весьма малой, достаточной для реального наблюдения интересующего нас явления при незначительных разностях в энергиях между изомерами, составляющих всего несколько десятков или сотен тысяч электрон-вольт, и отличии на несколько $\hbar$ единиц в моментах количества движения изомерных ядер.

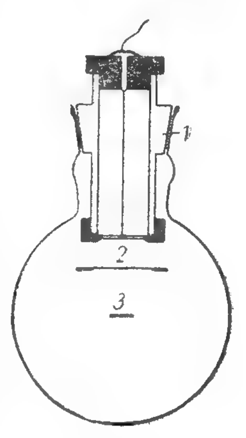
Открытие конверсионного излучения изомеров. Таким образом, к 1937 г. с большой достоверностью удалось экспериментально установить наличие ядерной изомерии в броме и в других элементах. Наряду с этим была предложена гипотеза для объяснения природы нового явления.

Согласно изложенной гипотезе о природе ядерной изомерии должен существовать новый вид радиоактивности — самостоятельный гамма-распад с относительно большим временем жизни, связанный с превращением одного ядерного изомера в другой.

Наряду с другими мы начали поиски этого нового вида радиоактивности. Было проведено исследование гамма- и бета-излучений, испускаемых всеми радиоактивными изотопами брома. Особое внимание было обращено на выяснение отношения числа гамма-квантов к общему числу бета-распадов. Предполагалось, что таким образом удастся найти ту часть гамма-лучей, которая связана с изомерными переходами метастабильных ядер.

Чтобы более точно определить указанное выше отношение числа гамма- и бета-лучей и исключить по возможности поглощение электронов в стенках счетчика Гейгер-Мюллера, радиоактивный препарат брома, счетчик Гейгер-Мюллера с отверстиями в цилиндрических стенках для пропу-

скания электронов и соответствующие фильтры были помещены внутрь откачиваемого сосуда. При проведении этих опытов (выполненных совместно с А. А. Юзефовичем) было замечено, что часть излучения, которая поглощалась в фильтре из алюминия, толщиною в 0.2 мм, была непостоянной и зависела от каких-то случайных причин. Возникло предположение, что причина неустойчивости результата связана с толщиной слоя целлофанного лака, покрывавшего радиоактивный препарат. Дальнейшие опыты подтвердили это предположение и показали, что при уменьшении слоя радиоактивного препарата и при покрытии его тонким слоем целлофана число отбросов от замеченного нового излучения составляло 25—30%, а иногда и 50% от общего числа измеряемых счетчиком электронов бета-распада. Однако отчетливого количественного результата на этом приборе получить не удалось, так как было выяснено, что при опускании различных фильтров на счетчик эффективность его счета заметно менялась. Это явление было связано с тем, что при перемещении фильтров изменялся потенциал цилиндра счетчика относительно нити.



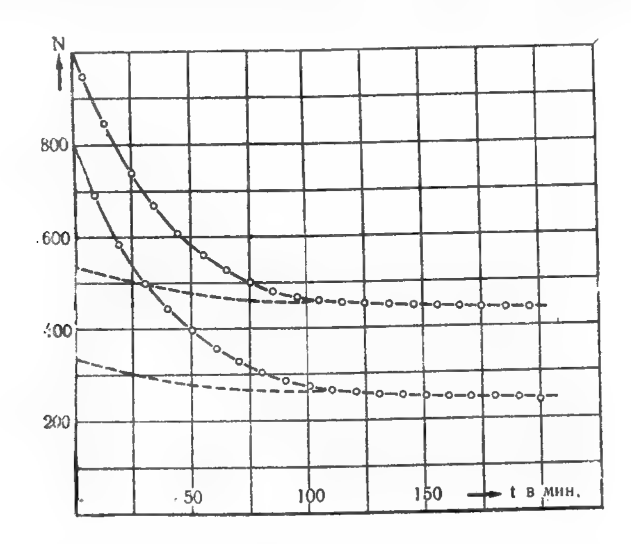
Фиг. 1
Прибор, с помощью которого изучалось мягкое электронное излучение брома.
1 — счетчик Гейгер-Мюллера; *2* — целлофанные фильтры; *3* — радиоактивный препарат

Был изготовлен новый прибор, в котором радиоактивный препарат был также помещен внутрь откачиваемого сосуда, но электроны направлялись вдоль нити счетчика Гейгер-Мюллера (фиг. 1).

Между препаратом и счетчиком помещались специальные фильтры для поглощения электронов. Контрольные опыты показали устойчивость и воспроизводимость результатов измерений на таком приборе. При уменьшении толщины фильтров нам удалось заметить, что фильтр толщиною всего лишь в 20 микрон целлофана целиком поглощает замеченное мягкое излучение, причем интенсивность этой «мягкой» электронной компоненты оказалась, примерно, равной интенсивности «жесткой», которая проходит через фильтр указанной выше толщины (фиг. 2). В процессе проведения опытов становилось очевидным, что мы имеем дело с новым интересным явлением.

19*

Опыты по отклонению мягкого излучения в магнитном поле показали, что новое излучение представляет собою поток отрицательных электронов. Измерения поглощения этого излучения в очень тонких фильтрах целлофана разной толщины установили, что энергия электронов может быть оценена приблизительно в 30 keV. Далее было выяснено,



Фиг. 2. Кривые распада радиоактивного брома: верхняя — без фильтра, нижняя — с фильтром целлофана толщиною в 20 микрон

что интенсивность мягкого излучения уменьшается с периодом полураспада в 4.2 часа. На основании всей совокупности фактов можно было заключить, что в случае радиоактивных превращений брома с периодом полураспада в 4.2 часа каждый бета-распад сопровождается испусканием мягкого электрона, обладающего указанной выше энергией.

В самой начальной стадии опытов, когда были замечены описанные аномалии в поглощении электронов, И. В. Курчатов выдвинул гипотезу о том, что наблюдаемое мягкое электронное излучение обусловлено процессами внутренней электронной конверсии в атомах брома и не связано с особенностями бета-распадов. Согласно этой гипотезе, изомеры брома, находящиеся в метастабильном состоянии, переходят в изомеры основного

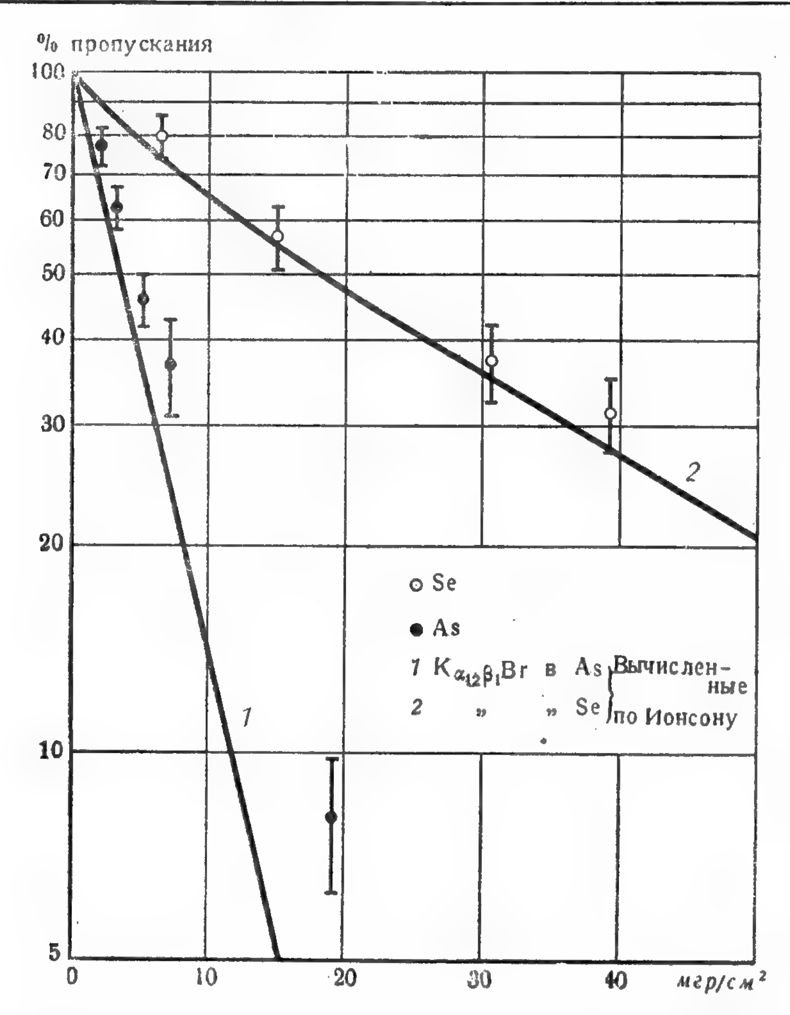
состояния главным образом путем процессов внутренней конверсии. Возбужденные изомеры брома передают свою энергию одному из ближайших к ядру электронов, который выбрасывается за пределы атома. Заметим, что при обычном испускании гамма-лучей возбужденными ядрами, получающимися при естественно радиоактивном распаде, наблюдались электроны конверсии, но они, как правило, составляли небольшую долю, порядка нескольких процентов, общего числа переходов. В случае же брома на основании проведенных наблюдений следовало предположить, что почти все превращения изомеров происходят путем процессов внутренней электронной конверсии.

В свете новых представлений становилась ясной причина тех затруднений, с которыми было связано обнаружение самостоятельной радиоактивности в виде гамма-лучей при высвечивании изомерных ядер, находящихся в метастабильном состоянии. Эти затруднения обусловлены самой природой явления — тем, что метастабильные ядра переходят в основное состояние главным образом путем процессов внутренней электронной конверсии. Новый вид гамма-радиоактивности обнаруживается, таким образом, по испусканию электронов конверсии. В дальнейшем были найдены ядерные изомеры, где конверсия оказалась частичной.

Идея о конверсионном происхождении наблюденного мягкого электронного излучения требовала экспериментального доказательства. Не исключалась возможность, что замеченное излучение было связано с особенностями формы бета-спектров радиоактивного брома.

Каким же образом можно было на опыте проверить гипотезу о конверсионном происхождении мягкого излучения? При вылете электронов из внутренних, ближайших к ядру слоев электронной оболочки атома их место заполняют наружные электроны. При этой перестройке электронной оболочки испускаются характеристические рентгеновы лучи, энергия которых равна разности энергии электронов на различных оболочках атома.

Согласно гипотезе о конверсионном происхождении мягкого электронного излучения, радиоактивный распад брома должен сопровождаться испусканием характеристических рентгеновых лучей. Для обнаружения этих рентгеновых лучей были поставлены опыты. Все электроны, испускаемые бромом, вплоть до энергии максимальной границы бета-спектра (~2 MeV) отклонялись магнитным полем. В качестве индикатора рентге-



Фиг. 3. Кривые поглощения рентгеновского излучения изомерного брома в мышьяке и селене (экспериментальные данные сопоставлены с вычисленными по формулам Ионсона)

новых квантов был использован счетчик Гейгер-Мюллера, наполненный аргоном, обладающий большой чувствительностью к их счету.

В первых же опытах, выполненных Л. И. Русиновым и А. А. Юзефовичем, было обнаружено при включенном магнитном поле достаточно

большое число импульсов в счетчике. Число импульсов сильно уменьшалось при фильтрации свинцом, толщиною в 1 мм, и, таким образом, можно было заключить, что оно относится к длинам волн рентгеновской области.

Длина волны найденного электромагнитного излучения вначале была оценена по поглощению в алюминиевом фильтре. Было установлено, что длина волны примерно равна 1 Å, что соответствовало Kα-линии рентгеновского спектра брома. Для более точного определения длины волны был применен очень чувствительный метод так называемых дифференциальных фильтров, который полностью подтвердил предварительные результаты и установил по поглощению в As и Se (фиг. 3), что радиоактивный источник брома испускает рентгеновы лучи, длина волны которых лежит в пределах 0.97773 и 1.04263 Å.

Далее было показано, что активность рентгеновского излучения уменьшается с периодом мягкого излучения (4.2 часа).

Рядом контрольных опытов была установлена невозможность приписать наблюденные рентгеновы лучи явлению захвата орбитального K-электрона ядрами брома.

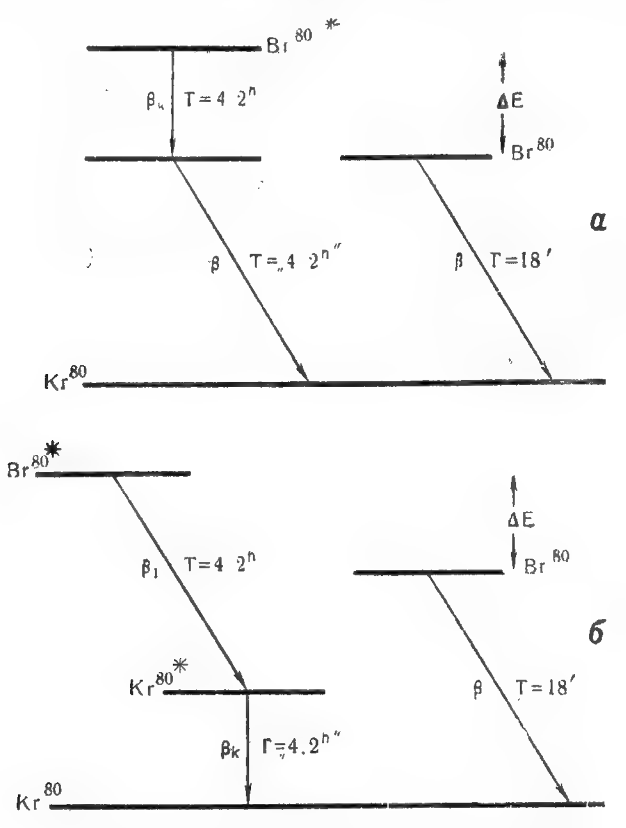
На основании всей совокупности экспериментов можно было с несомненностью сделать заключение, что наблюденное мягкое электронное излучение, испускаемое радиоактивным бромом, обусловлено процессами внутренней электронной конверсии, сопровождающими изомерные превращения брома.

К аналогичным выводам о распаде изомеров путем процессов конверсии пришел, на основании опытов с родием, Понтекорво. Полученные опытные данные об аномально больших коэффициентах конверсии при изомерных переходах находились в согласии с появившимися в то время теоретическими расчетами Хебба и Уленбека.

Процессы внутренней электронной конверсии, сопутствующие изомерным превращениям, открыли экспериментальные возможности для изучения существенных деталей ядерной изомерии и понимания природы явления.

Установление основных типов превращения изомеров. Принципиально возможны две схемы превращений при ядерной изомерии — схемы последовательного и независимого распада. Отличие между ними состоит в том, что при последовательном распаде изомер с бо́льшим запасом энергии, находящийся в метастабильном

состоянии, путем гамма-радиоактивности непосредственно превращается в другой, обладающий меньшей энергией, и находящийся в основном состоянии. При независимом же распаде два изомера распадаются путем



Фиг. 4. Схемы превращений при изомерии брома:

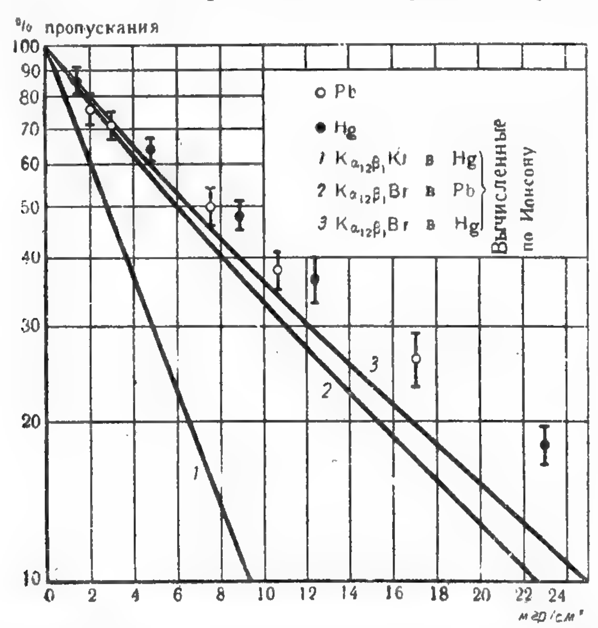
а — последовательная, *б* — независимая (β_k — конверсионный переход, β, β_1 — бета-распад)

бета-излучения на различные уровни соседних ядер, т. е. возбужденный изомер распадается, минуя основное состояние другого изомера.

Применительно к случаю брома схемы последовательного и независимого распада представлены на фиг. 4. Согласно опытам Боте — Гентнера, Снелла и др., удалось путем перекрестных ядерных реакций установить,

что изомерными являются ядра брома с атомным номером 80, которые распадаются с двумя периодами полураспада: 18 мин. и 4.2 часа. Существенно было выяснить, по какой из схем превращений происходит распад изомерного брома.

При обеих схемах превращения изомеров брома наблюдаемые периоды бета-распадов будут практически равны 18 мин. и 4.2 часа, а уменьшение интенсивности конверсионного излучения будет наблюдаться



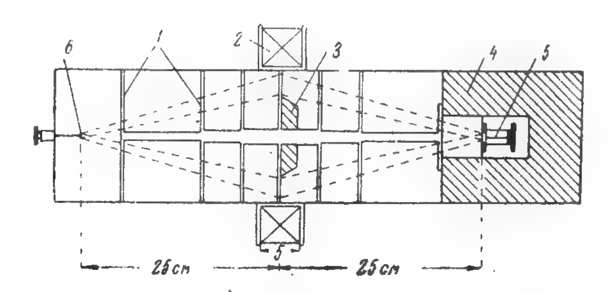
Фиг. 5. Кривые поглощения рентгеновского излучения изомерного брома в свинце и ртути (экспериментальные данные сопоставлены с вычисленными по формулам Ионсона)

с периодом 4.2 часа и, следовательно, по этим характеристикам две схемы не могут быть разделены.

Нами было обращено внимание на то обстоятельство, что при последовательном распаде процессы внутренней электронной конверсии должны происходить в атомах брома, а при независимом — в атомах криптона; поэтому в первом случае должно происходить характеристическое рентгеновское излучение брома, а во втором — криптона. Опыты по

поглощению рентгеновского излучения радиоактивного брома в Pb и Hg (фиг. 5), выполненные методом дифференциальных фильтров, показали, что испускаются характеристические рентгеновы лучи брома, а не криптона. Тем самым было доказано, что изомерные превращения в броме идут по схеме последовательного распада.

Определение разности энергии между изомерными ядрами Установление схемы переходов дало возмож-



Фиг. 6. Магнитный спектрограф для определения энергии электронов конверсии.

1 — алюминиевые диафрагмы; 2 — фокусирующая катушка; 3 — свинцовый фильтр; 4 — свинцовая защита; 5 – счетчик Гейгер-Мюллера; 6 — радиоактивный источник

ность определить разность энергии между изомерными ядрами брома. Ввиду того, что в случае изомерии брома имеет место последовательный распад изомеров, энергия возбуждения метастабильного ядра может быть установлена по энергии электронов внутренней конверсии.

Л. И. Русиновым и А. А. Юзефовичем на магнитном спектрографе с продольной фокусировкой (фиг. 6) был исследован энергетический спектр электронов конверсии. В результате измерений была получена кривая распределения электронов по энергиям (фиг. 7). Из этой кривой следует, что в спектре электронов имеются два максимума, соответствующие энергии $\sim$35 и 47 keV. Первый максимум следует отнести к ионизации ближайшей к ядру K-оболочки, а второй — к вырыванию электронов из L-слоя.

Из сопоставления полученных экспериментальных данных для энергии электронов конверсии на различных оболочках атомов с потенциалами

ионизации K- и L-электронов (13.6 и 1.8 keV), было установлено, что энергия возбуждения метастабильного ядра брома составляет $\sim$48 keV.

Определение разности угловых моментов изомеров. Для полного выяснения природы ядерной изомерии необходимо было показать наличие большой разности угловых моментов между изомерными ядрами. Интересно было бы также количественно определить величину этой разности и произвести общее сопоставление опытных данных с теорией.

Мы попытались выяснить вопрос об отличии угловых моментов между изомерами, используя процессы внутренней электронной конверсии. Из квантовой теории излучения следует, что изменение углового момента при переходе возбужденного ядра в основное состояние может быть определено по порядку мультипольности излучения, связанного с этим переходом. Наблюдаемые на опыте коэффициенты внутренней конверсии в сильной степени должны зависеть от порядка мультипольности излучения, которое в свою очередь определяется разностью угловых моментов превращающихся изомерных ядер. Следовательно, изменение разности угловых моментов излучающих ядер может быть выяснено путем сравнения опытных данных по конверсии с теоретическими расчетами, устанавливающими количественную связь между коэффициентами конверсии и порядком мультипольности излучения.

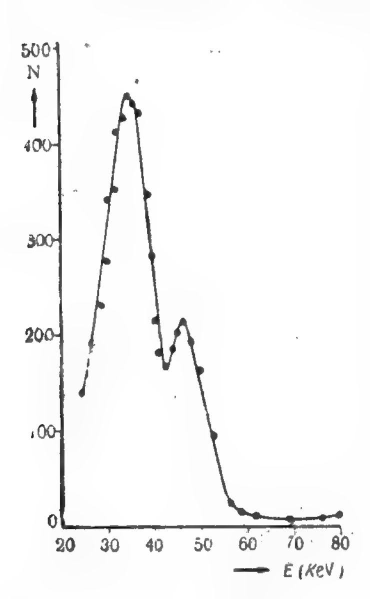
Рассчитываемые по теории значения вероятности процессов внутренней конверсии не требуют конкретизации модели излучающего ядра, сведения о котором пока весьма ограничены. Расчеты вероятностей электронной конверсии производятся по формулам волновой механики и могут быть выполнены в принципе с достаточной точностью.

Первые вычисления коэффициентов конверсии на K-оболочках атомов при излучениях высокой мультипольности принадлежат Хеббу и Уленбеку, а также Данкову и Моррисону. Они проведены для излучений электрического характера по нерелятивистским формулам и имеют область применимости для ядер с атомными номерами до 40 и энергией возбуждения менее 200 keV.

Для экспериментального определения коэффициентов конверсии при изомерных превращениях брома был снят общий электронный спектр, испускаемый изомером брома, имеющим константу распада 4.2 часа. На кривой было обнаружено два отчетливых максимума, соответствующих

конверсии на K- и L-оболочках атомов брома, и непрерывное распределение электронов основного бета-спектра. Число электронов в конверсионных максимумах соответствует числу превращений метастабильного ядра путем процессов внутренней электронной конверсии.

Число электронов бета-спектра, соответствующего превращению ядер Br^{80} в Kr^{80}, равно общему числу переходов метастабильного ядра Br^{80*} в основное состояние Br^{80}, происшедших как путем процессов внутренней конверсии, так и неконвертированным путем. Очевидно, по этим данным может быть определено, какая доля из общего числа переходов метастабильного ядра происходит путем внутренней конверсии. На основании полученной кривой удалось установить, что отношение числа актов конверсии, имеющих место на K-слое атомов брома, к числу переходов, происходящих путем неконвертированного гамма-излучения, составляет более 9.3. Из сопоставления полученного на опыте результата с теоретическими формулами следовало, что изменение углового момента в случае изомерного перехода брома больше или равно 2 $\hbar$ единицам.



Фиг. 7. Энергетический спектр электронов внутренней конверсии, испускаемых изомерным бромом

Мы попытались использовать возможность определения разности угловых моментов изомерных ядер из тех же экспериментальных данных по величине отношения коэффициентов конверсии на K- и L-оболочках атомов брома. Опыты показали, что относительная вероятность конверсии на L-оболочке по сравнению с K-слоем составляет ~ 0.5 (фиг. 7).

Принимая это значение и используя полученные Г. С. Завелевичем теоретические данные для относительных коэффициентов конверсии, можно заключить, что метастабильное ядро брома испускает октупольное излучение с изменением углового момента на 3 $\hbar$ единицы.

Полученные данные об отличии угловых моментов изомеров брома

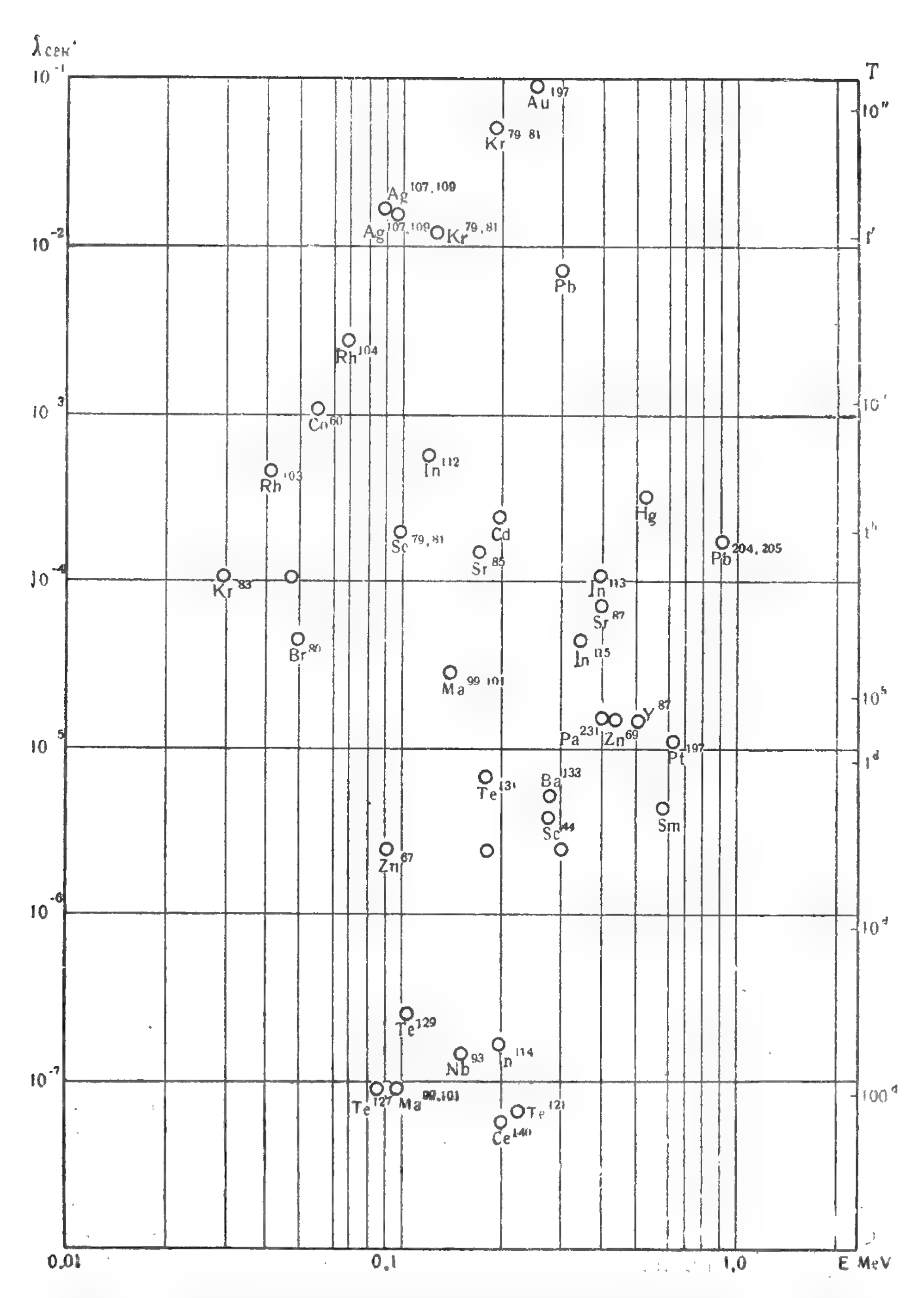
подтвердили теоретические представления о том, что причина метастабильности изомеров определяется большой разностью угловых моментов их атомных ядер.

Современное состояние вопроса о ядерной изомерии. После открытия ядерной изомерии брома это явление было установлено у значительного числа других атомных ядер. Изомерные состояния встречаются у многих элементов периодической системы, и число открытых пар ядерных изомеров приближается к 70.

Энергии возбуждения изомеров, измеренные в большинстве случаев аналогично брому по электронам конверсии, составляют от ~50 keV (Br^{80} и Kr^{83}) до ~500 keV (Zn^{69} и др.); наблюденные периоды полураспадов варьируют от ~7.5 сек. (Au^{197}) до ~5 месяцев (Ce^{140}).

Все ядерные изомеры, для которых имеются определенные данные для времени жизни метастабильного изомерного состояния и энергии возбуждения изомера, представлены на фиг. 8. В ближайшем будущем следует ожидать дальнейшего увеличения числа ядерных изомеров, в частности открытия изомеров с очень малым и очень большим временем жизни. Обычными методами искусственной радиоактивности очень трудно обнаружить ядерную изомерию в том случае, когда время жизни метастабильного состояния меньше 10 сек. С другой стороны, затруднительно и обнаружение долго живущих изомеров. В то же время по природе явления следует ожидать, что ядра-изомеры могут иметь весьма широкий диапазон для времени жизни метастабильных состояний — от 10^{-13} сек. до любых больших промежутков времени.

И. В. Курчатовым и сотрудниками Радиевого института Академии Наук СССР были проведены опыты по обнаружению короткоживущей изомерии у ряда элементов. Обнаружения изомерии производились по действию на фотографическую пленку электронов конверсии, которые могли излучаться при изомерных переходах ядрами, возникающими при захвате медленных нейтронов. Постановка опыта обеспечивала обнаружение электронов конверсии в тех случаях, если время жизни метастабильного состояния очень мало. При облучении гадолиния нейтронами было замечено почернение фотографической пленки. Контрольные опыты показали, что оно вызвано электронами внутренней конверсии с энергией около 100 keV. Ферми позднее показал, что время жизни изомерного состояния гадолиния около 10^{-3} сек.



Фиг. 8. График изомерных ядер, для которых получены данные о времени жизни метастабильного состояния и энергии возбуждения изомерного ядра.

Можно предположить, что среди обычных элементов существует ядерная изомерия с периодом полураспада порядка 10^9 лет, сравнимым со временем жизни Земли. Л. И. Русинов и Я. М. Игельницкий пытались экспериментально обнаружить у ряда элементов такую изомерию, но не пришли к определенным результатам.

Время жизни ядер по отношению к тому или иному виду ядерного излучения характеризует процессы, происходящие в ядре. Изучение зависимости между этими временами жизни и значениями энергий излучений сыграло большую роль в физике атомного ядра. Анализ вопроса о связи периодов альфа- и бета-распадов с энергией испускаемых частиц привел ко многим важным представлениям о природе происходящих в ядре процессов. Соображения Бора о времени испускания гамма-лучей при захвате нейтронов создали очень важные для теории ядер представления о наличии при ядерных реакциях промежуточного состояния ядра.

Новое явление гамма-радиоактивности при изомерных переходах характеризует продолжительность жизни ядер по отношению к запрещенному гамма-излучению. Поэтому мы вправе ожидать, что ядерная изомерия даст возможность глубже понять происходящие при электромагнитном излучении ядра процессы.

В настоящее время имеются попытки теоретически установить количественную связь между временем жизни изомеров, энергией возбуждения метастабильного состояния и разностью угловых моментов изомерных ядер. Подобные расчеты были сделаны Вайцзекером, Бете, Ландау, Сегре, Хеббом и Уленбеком, Лоуеном, Бертлоу и другими. Авторы исходили из различных моделей ядра как излучающей системы, рассматривая излучения ядер как результат колебания протона, а в ряде случаев альфа-частицы в самосогласованном поле остальных частиц ядра, изучая радиацию от колебаний равномерно заряженной жидкой капли, а также трактуя излучение ядерных мультиполей по общим законам квантовой электродинамики.

Подсчет времени жизни изомерных ядер по разным теориям дает очень большое расхождение между собой (до 10^{10} раз) при предположении одной и той же разности энергии и одинакового отличия угловых моментов изомеров.

Полное сравнение опытных данных с теориями ядерной изомерии с целью выяснения, какая из теорий имеет наиболее близкое отношение

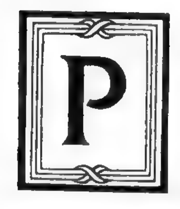
к действительности, может быть проведено только для очень небольшого числа случаев. Самым сложным является вопрос о точном определении разности угловых моментов между изомерами. Это определение угловых моментов может быть выполнено тем же путем, как и для брома, — путем исследования коэффициентов внутренней конверсии или же по корреляции гамма-квантов и электронов конверсии. Современные интенсивные источники нейтронов дают возможность получить искусственно радиоактивные вещества в весомых количествах. В связи с этим возникает возможность определения угловых моментов изомеров прямыми путями, известными в атомной физике — изучением сверхтонкой структуры оптических спектров и методами отклонения молекулярных пучков в неоднородном магнитном поле.

Дальнейшее изучение ядер изомеров и гамма-радиоактивности при ядерной изомерии, несомненно, расширит наши представления о структуре атомного ядра.

Академик

А. Ф. Иоффе

ЭЛЕКТРОННЫЕ ПРОВОДНИКИ ТОКА



Разнообразие электрических свойств различных тел чрезвычайно велико. Удельная проводимость металлов измеряется сотнями тысяч обратных омов, тогда как удельное сопротивление изоляторов достигает 10^{15}—10^{20} ом·см. Электропроводности тел отличаются больше чем на 20 порядков, в то время как теплопроводность наиболее далеко отстоящих по своей структуре твердых тел различается не более чем на 3—4 порядка.

По величине удельного сопротивления проводники и изоляторы занимают крайние места, между ними находится обширная область полупроводников.

Различна и физическая природа зарядов в разных веществах. В одних электрический заряд неразрывно связан с атомом или группой атомов — это ионные проводники тока или электролиты; в других носителями тока оказываются лишенные индивидуальности электроны; перемещение одних электронов на место других не вносит химических изменений. Замкнутая цепь из таких проводников не может создавать при постоянной температуре электродвижущих сил. Наблюдается и смешанная проводимость: ионы и электроны образуют ток в любых соотношениях.

Среди ионных и среди электронных проводников имеются и хорошо проводящие материалы (металлы и расплавленные соли, через которые

в крупных технических установках проходят тысячи ампер) и изоляторы (например, электронная двуокись титана и ионная каменная соль). Как в ионных, так и в электронных веществах большую, часто решающую роль играют посторонние примеси или небольшие отступления от точного стехиометрического состава.

Меньшее значение имеют, как оказалось, механические искажения структуры.

Квантовые уровни электронов в твердом теле. Несмотря на столь большие количественные и качественные различия, современная волновая механика объединяет все тела одной общей теорией электрических свойств. Основу этой теории составляют условия закрепления электронов в теле — их энергетические уровни.

Со времени первой теории Бора мы знаем, что в атоме электроны могут занимать только вполне определенные стационарные квантовые уровни энергии и не могут обладать промежуточными ее значениями. Когда несколько одинаковых атомов соединяются в молекулу или когда громадное число N атомов образует твердое тело, одинаковые раньше уровни расщепляются на такое же число N близких друг к другу, но различных между собою уровней. Во всем теле нет тогда электронов, обладающих одинаковыми квантовыми числами, как их не может быть и в отдельном атоме. Так как число атомов, а следовательно и число одинаковых электронов, даже в небольшом кристалле чрезвычайно велико, порядка 10^{15}—10^{20}, то энергии расщепленных уровней оказываются очень близкими друг к другу. С достаточным приближением мы можем их рассматривать как сплошную полосу уровней.

Основные типы проводников электричества. Как в отдельном атоме, так и в кристалле мы, строго говоря, должны рассматривать все положительные ядра тела и всю совокупность взаимодействующих с ними электронов как одну квантовую систему. Именно поэтому во всем теле не должно быть двух электронов с одинаковыми квантовыми числами. Однако энергии глубоких электронов так слабо искажены соседними атомами, что каждый из них можно считать связанным только со своим ядром.

Наоборот, энергия валентных электронов настолько изменена окружающими полями, что их приходится объединять в коллектив, заполняющий полосу квантовых уровней всего кристалла.

Как в отдельном атоме, так и в кристалле, кроме уровней наинизшей энергии, нормально заполненных электронами, существуют незаполненные квантовые уровни с повышенной энергией — уровни возбуждения, которые в кристалле, так же как и основные уровни, расщепляются в полосы.

Зная систему энергетических уровней валентных электронов, можно предвидеть электрические свойства тела: если число уровней в полосе равно числу валентных электронов, то все уровни полосы заняты и ни один из электронов не может изменить своего состояния внутри полосы; замена же одного электрона другим никаких внешних проявлений не создает. Электрическое поле в таком теле не изменяет движения электронов и не создает тока. В таком теле не может быть электронной проводимости; это — изолятор или электролит.

Однако, кроме основных уровней, существуют и уровни возбуждения с большей энергией, на которые электрон может перейти, если ему сообщить необходимую для этого энергию. Источником этой энергии может быть тепловое движение, поглощенный фотон или сильное электрическое поле. Чем меньше разность энергий между нормальной и возбужденной полосой уровней или, другими словами, чем у́же ширина запретной энергетической зоны, их разделяющей, тем меньше энергии потребуется на ее преодоление. Ширина запретной зоны определяет число электронов, которое может перейти за единицу времени в полосу возбуждения при данной температуре, красную границу спектра, создающего фотопроводимость или напряжение поля, вызывающего повышенную проводимость. Это — область электронных полупроводников.

Если нормальная и возбужденная полосы уровней перекрываются или если в нормальной полосе остается значительное число уровней, не замещенных валентными электронами, переход из одного квантового состояния в другое становится возможным под влиянием любого, самого слабого электрического поля и без помощи тепловой энергии. Такие тела мы называем металлами.

Как видим, электрическое поведение тела определяется не числом электронов, которое одинаково в металле и изоляторе, а возможностью изменения квантового состояния электрона при внешнем воздействии.

М е х а н и з м т о к а. Когда часть электронов переходит из сплошь заполненной полосы в свободную полосу возбуждения, то здесь они

20*

получают возможность свободно изменять свое состояние. Если, например, без внешнего поля одинаковое число электронов двигалось слева направо и справа налево, то под воздействием поля появится статистическое преобладание электронов, движущихся по направлению действующих на них сил — пойдет электрический ток, создающий вокруг себя магнитное поле и вызывающий переход затраченной работы в тепло.

Переход электронов в полосу возбуждения видоизменяет и основную полосу, освобождая в ней ряд квантовых уровней и устраняя причину, не позволявшую совокупности электронов следовать внешним воздействиям. Теперь электроны нормальной полосы также получают возможность участвовать в прохождении электрического тока, переходя на свободные уровни и освобождая свои места для других электронов.

Не следует забывать, что электроны в кристалле движутся в сложном электрическом поле и поэтому не могут рассматриваться как «свободные», подобно электронам в вакууме.

Волновая механика показывает, что в строго периодическом поле неискаженного кристалла к электрону могут быть применены законы движения свободных зарядов, с той разницей, что место универсальной массы электрона $9 \cdot 10^{-28}$ г занимает «эффективная масса» — величина, зависящая от строения кристалла и положения электрона в системе квантовых уровней данной полосы.

Обычно эффективная масса электрона в свободной полосе по порядку величины близка к массе свободного электрона. Свободные же места, образовавшиеся вследствие ухода электронов наибольших энергий из нормальной полосы, эквивалентны электронам не только с иной массой, но и с положительным зарядом. Для краткости их часто называют «положительными дырками», а вызванную ими проводимость — «дырочным» механизмом проводимости.

Электрический ток складывается из потока отрицательных электронов в одну сторону и перемещения положительных дырок в противоположную. В поперечном к току магнитном поле электроны и дырки создают противоположные электрические поля; по их направлению можно судить о том или ином механизме тока.

Другой способ, позволяющий узнать знак носителей тока, — это определение термоэлектродвижущих сил, возникающих в электронном полупроводнике. В нагретом конце концентрация носителей тока возра-

стает, и они диффундируют к холодному концу, заряжая его своим зарядом и оставляя противоположный заряд на горячем конце. Так как термоэлектродвижущие силы в полупроводниках гораздо больше, чем в металлах, то указанное правило применимо и к цепи, составленной из интересующего нас полупроводника и любого металла.

Примеси в полупроводниках. В идеальном кристалле число «дырок» в нормальной полосе должно равняться числу электронов в полосе возбуждения, и поэтому оба механизма тока должны существовать одновременно, в значительной степени уравновешивая друг друга. В реальных кристаллах, однако, всегда имеется хотя бы количественно ничтожная примесь посторонних веществ, и самое соотношение числа атомов, из которых построен кристалл, несколько отличается от точного стехиометрического состава.

Эти небольшие отступления оказывают часто решающее влияние на свойства как ионных, так и электронных полупроводников.

Посторонние, излишние или недостающие, атомы дополняют систему энергетических уровней. Часто уровни этих искажений оказываются внутри запретной зоны. Они могут или захватывать электроны заполненной полосы, освобождая в них «дырки», или же передавать свои валентные электроны в свободную полосу.

Несмотря на то, что по самому смыслу слова «примесь» число добавочных уровней всегда мало по сравнению с электронами основной решетки кристалла, создаваемая ими добавочная электропроводность может далеко перевешивать электропроводность чистого кристалла.

Действительно, если ширина запретной зоны составляет ΔE, причем ΔE значительно больше средней тепловой энергии kT при данной абсолютной температуре T, то из всех электронов только те немногие, которые обладают энергией, превышающей ΔE, перейдут в свободную полосу. В то же время электроны свободной полосы, обладающие повышенной энергией, рекомбинируют, отдавая избыточную энергию кристаллу. В состоянии статистического равновесия концентрация n электронов, находящихся в полосе возбуждения, определится выражением

$$n = aNe^{-\frac{\Delta E}{kT}}, \text{ или } n = aN^{\frac{1}{2}}e^{-\frac{\Delta E}{2kT}}, \qquad (1)$$

где N — общее число электронных уровней.

Множитель $e^{-\frac{\Delta E}{kT}}$ в высокой степени зависит от значения ΔE, когда $\Delta E \gg kT$. Например, при $kT = 0.03$ эл.-в., что соответствует температуре, немного превышающей комнатную, и при

$$\Delta E = 0.15 \text{ эл.-в.} = 5\,kT \qquad e^{-\frac{\Delta E}{kT}} \approx 10^{-3}$$

$$\Delta E = 0.3 \text{ эл.-в.} = 10\,kT \qquad e^{-\frac{\Delta E}{kT}} \approx 5\cdot 10^{-5}$$

$$\Delta E = 0.45 \text{ эл.-в.} = 15\,kT \qquad e^{-\frac{\Delta E}{kT}} \approx 3\cdot 10^{-7}$$

$$\Delta E = 0.6 \text{ эл.-в.} = 20\,kT \qquad e^{-\frac{\Delta E}{kT}} \approx 2\cdot 10^{-9}$$

$$\Delta E = 0.9 \text{ эл.-в.} = 30\,kT \qquad e^{-\frac{\Delta E}{kT}} \approx 10^{-13}$$

$$\Delta E = 1.5 \text{ эл.-в.} = 50\,kT \qquad e^{-\frac{\Delta E}{kT}} \approx 3\cdot 10^{-21}$$

Таким образом, например, для уровней примеси, отличающихся от свободной зоны на 0.3 эл.-в., множитель $e^{-\frac{\Delta E}{kT}}$ в 10^8 раз больше, чем для нормальной полосы, отделенной запретной зоной в 0.9 эл.-в. Если даже примесные уровни составляют только 0.01%, или 10^{-4} от числа основных уровней, то они все же создадут еще в 10^4 раз больше электронов. Понятно, что в случаях, подобных приведенному численному примеру, концентрация свободных электронов определяется почти исключительно уровнями примеси.

Но когда мы перейдем в том же примере от температуры $T \approx 330°$K, при которой $kT = 0.03$ эл.-в., к более высокой температуре $T \approx 1.300°$K, при которой $kT = 0.1$ эл.-в., то получим для уровней примеси $e^{-\frac{\Delta E}{kT}} = 5\cdot 10^{-2}$, а для основной полосы $e^{-\frac{\Delta E}{kT}} = 3\cdot 10^{-4}$. Та же примесь

дает теперь в 30 раз меньше свободных электронов, чем основная решетка кристалла.

Как показали В. П. Жузе и Б. В. Курчатов, число свободных электронов (или дырок) примесного полупроводника при любой температуре складывается из n_1 электронов, создаваемых основной решеткой

$$n_1 = aN_0^{\frac{1}{2}} e^{-\frac{\Delta E_0}{2kT}},$$

где N_0 — число атомов в 1 см^3 кристалла, и из n_2 электронов, создаваемых примесными уровнями

$$n_2 = aN^{\frac{1}{2}} e^{-\frac{\Delta E}{2kT}},$$

$$n = n_1 + n_2 = aN_0^{\frac{1}{2}} e^{-\frac{\Delta E_0}{2kT}} + aN^{\frac{1}{2}} e^{-\frac{\Delta E}{2kT}}.$$

В случае исследованной ими закиси меди Cu_2O $\Delta E_0 = 1.44$ эл.-в., $\Delta E \approx 0.6$ эл.-в., а примесь кислорода 0—1%.

Формулы, подобные (1), удобно изображать графически, предварительно логарифмируя их. Тогда

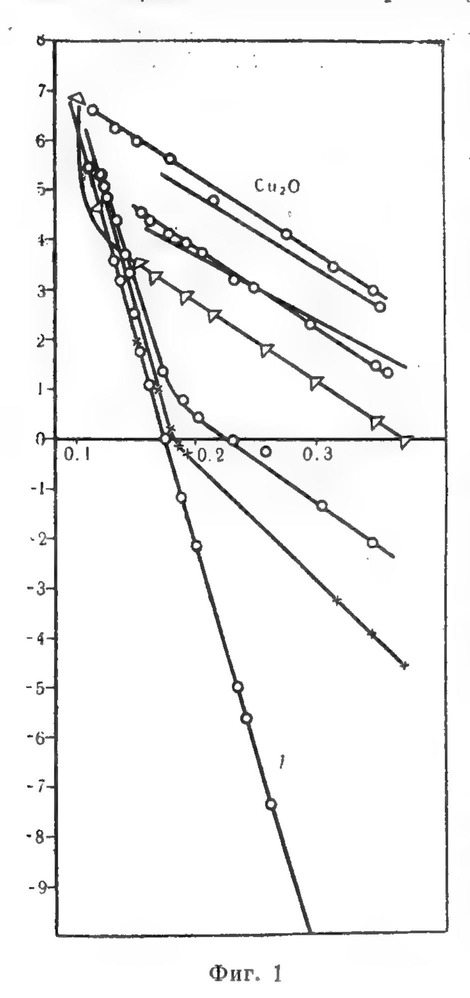
$$\log \sigma = \log a - \frac{\Delta E}{k} \frac{1}{T}, \qquad (1a)$$

где $\log a$ и $-\frac{\Delta E}{k}$ —постоянные. В системе координат, где по оси абсцисс нанесены значения $1/T$, а по оси ординат $\log \sigma$ (фиг. 1), уравнение (1) дает прямую линию, пересекающую ось ординат (при $1/T = 0$ или $T = \infty$) на высоте $\log \sigma = \log a$.

Тангенс угла, образуемого прямой с осью абсцисс, определяет постоянную $-\frac{\Delta E}{k}$, где k — универсальная постоянная, равная $1.38 \cdot 10^{-16}$ эрг/град.

Таким образом, график фиг. 1 дает возможность определить энергетический барьер, необходимый для перевода электрона в свободное состояние. Прямая 1 представляет зависимость $\log \sigma$ от обратной абсолютной температуры $1/T$ для основной решетки, а остальные прямые определяют добавочную электропроводность, созданную избытком кислорода.

Примеси по своей химической природе могут обладать свойством легко отщеплять свои электроны (таково большинство металлов); тогда они под влиянием теплового движения отдают свои электроны свободной полосе и создают электронный механизм проводимости.



Фиг. 1

Электроотрицательные примеси (например, кислород или сера), наоборот, захватывают электроны из нормальной заполненной полосы, освобождая в ней свободные уровни и создавая дырочный механизм тока.

Измерение величин, определяющих свойства зарядов. Когда носителями тока являются либо одни только электроны, либо одни дырки, концентрацию их можно определить, измерив электродвижущую силу E, возникающую при прохождении тока I через полупроводник, помещенный в поперечное магнитное поле H (эффект Холла)

$$E = RBIH, \tag{2}$$

где B — ширина пластинки, через которую проходит ток I. Коэффициент R определяет концентрацию n носителей тока

$$R = \frac{1}{en}. \tag{3}$$

Электропроводность σ зависит не только от концентрации носителей n, но и от их подвижности u, т. е. средней скорости, с которой перемещаются электроны в электрическом поле в 1 в/см:

$$\sigma = neu, \tag{4}$$

где e — заряд носителя тока, равный $\pm 1.6 \cdot 10^{-19}$ кулона. Измерив R и σ, мы получаем

$$u = \frac{1}{e} R\sigma. \tag{5}$$

Значительно сложнее обстоит дело, когда в прохождении тока одновременно участвуют и электроны и дырки. Однако и здесь, измеряя на одном и том же образце σ, R, термоэлектродвижущую силу и изменение сопротивления в продольном магнитном поле, можно определить число носителей тока. Задача облегчается, когда нам удается введением примесей обеспечить либо один, либо другой знак носителей тока.

Впервые изменение знака при химической обработке полупроводника Tl_2S было установлено Б. М. Гохбергом и М. С. Соминским. Ю. П. Маслаковец и Б. Т. Коломиец, вводя избыток таллия или серы, по произволу меняли знак носителей тока. Аналогичные результаты были получены Ю. П. Маслаковцем и Ю. А. Дунаевым для PbS и др.

Таким образом, удается измерить значения n_+, n_-, u_+, u_-, m_+, m_- для ряда полупроводников. Здесь знаки (+) относятся к «дыркам», (—) к электронам, а m — эффективная масса.

Фотоэлектрические свойства полупроводников. Переход электронов в свободную полосу может совершаться не только под влиянием теплового движения, но и вследствие поглощения света. А. В. Иоффе и А. Ф. Иоффе показали, что добавочное число электронов, создаваемое светом независимо от длины волны, пропорционально числу поглощенных фотонов. Диффузия зарядов из освещенного участка в соседние и перенос их током создают целый ряд на первый взгляд неожиданных явлений, которые, однако, удовлетворительно могли быть объяснены с этой точки зрения.

Так, в неоднородно освещенном кристалле возникают фотоэлектродвижущие силы; добавочная фотопроводимость и ее спектральное распределение зависят от величины и знака электрического поля, а в определенных условиях можно наблюдать вместо фотопроводимости уменьшение электропроводности при освещении.

И. К. Кикоин и М. М. Носков обнаружили в охлажденной жидким воздухом закиси меди, помещенной в магнитное поле, появление электродвижущих сил, превышающих 10 в, и провели полное эксперименталь-

ное и теоретическое исследование открытого ими фотомагнитного эффекта.

Сильные поля. А. В. Иоффе и А. Ф. Иоффе исследовали также повышение электропроводности под влиянием сильных электрических полей для большого числа различных полупроводников. Установлено было, что эффект этот наступает, только начиная с определенного напряжения электрического поля тем большего, чем шире запретная зона уровней. Было показано, что эффект сильного поля сводится к переводу электронов в свободную полосу уровней и что сильные поля не изменяют подвижности носителей тока. Определена зависимость добавочной проводимости от силы поля и от температуры, выяснено, что механизм повышения электропроводности в сильных полях не связан с механизмом электрического пробоя диэлектриков.

Контактные явления. Контакт полупроводника с металлом или с другим полупроводником характеризуется рядом своеобразных явлений, нашедших широкое применение в технике (твердые выпрямители и фотоэлементы). А. В. Иоффе исследовала эти явления экспериментально, а Б. И. Давыдов разработал их теорию.

А. В. Иоффе показала, в согласии с данными теории диффузионного равновесия электронов, что в полупроводниках с электронным механизмом тока возникает значительное добавочное сопротивление в слое, граничащем с проводником более высокого контактного потенциала (т. е. с проводником, требующим более высокой работы выхода из него зарядов). В «дырочных» полупроводниках, наоборот, добавочное сопротивление возникает в слое, граничащем с материалом меньшего контактного потенциала, чем полупроводник. Это правило подтвердилось на большом и разнообразном материале, в различных температурных условиях и в широком интервале электропроводностей.

Проводимость пограничных слоев зависит от направления проходящего через границу тока. Однако, как показала А. В. Иоффе, этот «выпрямительный» эффект гораздо резче выражен на контакте двух полупроводников с различным механизмом проводимости, чем в слое полупроводника, граничащего с металлом.

Незаконченные исследования. Наряду с электропроводностью изучалась и теплопроводность полупроводников. И здесь обнаружено было значительное влияние примесей, например снижение

теплопроводности вдвое при введении 1—2% примеси. Измерен ход теплопроводности с температурой, зависимость ее от состава и т. п. Работа в этом направлении еще далеко не закончена.

Интересные результаты получены по вопросу о длительности существования заряда в полосе проводимости и о вероятности рекомбинации свободных электронов в зависимости от температуры. Эти исследования не привели еще к законченной картине явления.

Исследуются также процессы образования и исчезновения сопротивления в пограничных слоях и распределение электропроводности и объемных зарядов в этих слоях.

Технические выходы. Современная электропромышленность широко использует полупроводники. По стоимости продукции промышленность полупроводников в США не уступает производству электрических машин.

Важнейшим видом продукции являются твердые выпрямители из закиси меди и селена. П. В. Шаравский и А. З. Левинзон полностью разработали технологию их производства и передали свой опыт заводам. Б. В. Курчатов, Ю. А. Дунаев и А. З. Левинзон создали, кроме того, новый тип выпрямителя из сернистой меди, выпрямляющий в 100 раз большие токи при одинаковой площади контакта. Эти выпрямители используются в сварочных аппаратах.

Свыше 15 лет известны также твердые фотоэлементы (сначала из закиси меди, а потом из селена), которые дают при освещении электрический ток без внешней электродвижущей силы. Б. Т. Коломиец и Ю. П. Маслаковец разработали новый фотоэлемент из сернистого таллия, который дает при освещении в 1 люмен ток от 6 до 10 тыс. микроампер, вместо 400—500 микроампер, получаемых от селеновых фотоэлементов. Такое увеличение фоточувствительности позволило использовать их в звуковом кино вместо вакуумных фотоэлементов, требующих источников высокого напряжения. Фотоэлементы Коломийца дают гораздо более чистый звук и упрощают киноаппаратуру.

Изучение спектрального распределения внутреннего фотоэффекта обнаружило ряд полупроводников, чувствительных в далекой инфракрасной части (PbS, Bi_2Se_3 и др.).

Б. Т. Коломиец и А. З. Левинзон разработали также несколько типов «термисторов» — приборов, повышающих после включения силу тока с

заданной скоростью и служащих для постепенного включения электрических устройств или играющих роль реле времени.

А. Н. Арсеньева обнаружила в ряде полупроводников резкое и обратимое изменение сопротивления при деформации. Этот «тензометрический» эффект может служить как для регистрации быстро меняющихся деформаций, так и для телефонных аппаратов.

Наконец, Ю. П. Маслаковец, А. Н. Воронин, Ю. А. Дунаев, Е. Д. Девяткова совместно с НИИ № 627 Министерства электропромышленности добились значительных успехов в задаче прямого превращения тепловой энергии в электрическую с помощью термобатарей. Коэффициент полезного действия удалось поднять до нескольких процентов и осуществить технически оформленную продукцию термоэлектрических генераторов.

Полупроводники большой удельной проводимости. Весь наш предыдущий опыт основан на изучении полупроводников с большим сопротивлением, что очень упрощает задачу: большая ширина запрещенной зоны делает невероятным появление примесей, создающих одновременно проводимости электронного типа в свободной полосе уровней и проводимость дырочного типа среди почти заполненной полосы. Рост числа свободных электронов с повышением температуры настолько преобладает над такими факторами, как подвижность, что значения электропроводности, постоянной Холла, термоэлектродвижущих сил почти целиком зависят от одной постоянной — расстояния между уровнями источников зарядов и соответственной полосой кристалла.

Такие эффекты, как фотопроводимость, фотоэлектродвижущие силы, легко наблюдаются в полупроводниках большого удельного сопротивления.

Однако, как для более полного понимания электрических свойств, так и для технического использования особый интерес представляют полупроводники с очень малым удельным сопротивлением. Наши исследования показали, что к ним следует также отнести сплавы многих металлов, по составу приближающиеся к стехиометрическим соотношениям.

К. С. Шифрин разработал теорию полупроводников, обладающих малым удельным сопротивлением вследствие ли узости запретной энергетической зоны, или благодаря присутствию большого числа до-

бавочных уровней примеси вблизи заполненной или свободной зоны. Здесь наблюдается ряд особенностей, отличающих их от обычных полупроводников.

Тепловое движение переводит в свободную зону такое количество электронов, что они сплошь заполняют все уровни с энергиями ниже 0.1—0.15 эл.-в. Поэтому дальнейшее повышение концентрации электронов в свободной зоне требует не только работы, необходимой для достижения ее нижнего края, но еще дополнительной энергии в 0.1—0.15 эл.-в. для того, чтобы электрон мог попасть на незамещенный уровень. Это обстоятельство значительно видоизменяет как температурный ход проводимости, так и другие характеристики полупроводника.

Из числа полупроводников малого удельного сопротивления мы изучили сернистый свинец PbS. Избыток свинца создает электронную проводимость, избыток серы — дырочную. Величина удельной электропроводности в зависимости от избытка свинца или серы достигает нескольких тысяч обратных омов на 1 см, в то время как стехиометрический PbS обладает проводимостью порядка нескольких обратных омов.

Свинец или сера, введенные в PbS, как показал Ю. А. Дунаев, частично растворяются в нем, являясь источником добавочных носителей тока, частично же выпадают в виде коллоидных крупинок, не принимая участия в переносе тока. С повышением температуры предельная концентрация растворенного Pb или S повышается. При быстром охлаждении свинец или сера остаются в атомном виде и только постепенно, и тем скорее, чем выше температура, избыток атомов коагулирует, приближая образец к состоянию равновесия.

Таким образом, можно получать PbS с избыточной концентрацией растворенного свинца или серы. Такие «закаленные» образцы переходят в нормальное состояние путем отжига.

О концентрации n электронов или дырок можно судить по постоянной R эффекта Холла и по величине термоэлектродвижущей силы α, вызываемой разностью температур в 1° С.

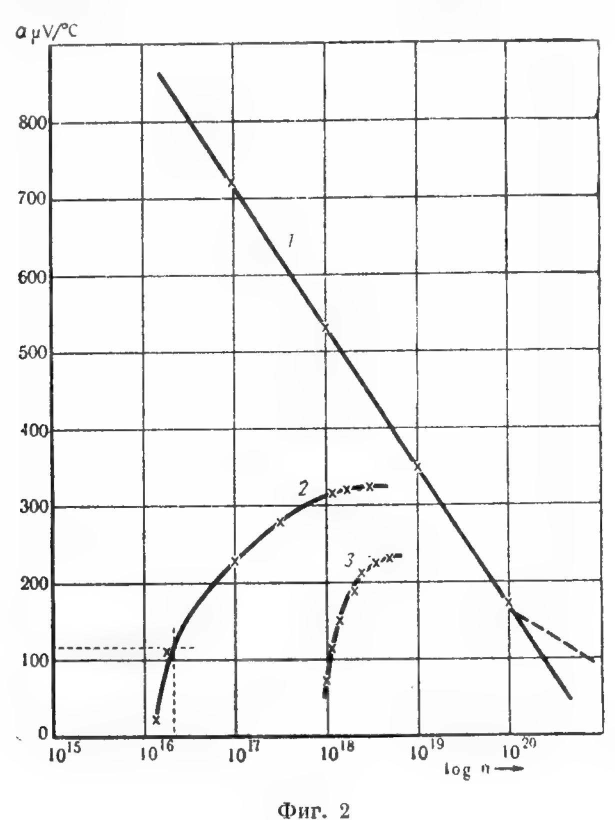
Постоянная эффекта Холла R в том случае, когда ток переносится либо одними электронами, либо одними дырками (что обычно и имеет место), равна, как мы видели,

$$R = \pm \frac{1}{en},$$

где e — заряд электрона. Термоэлектродвижущая же сила α в этих случаях

$$\alpha = \pm \frac{k}{e}\left(2 + \log \frac{(2\pi mkT)^{3/2}}{nh^3}\right) \qquad (6)$$

также однозначно определяется концентрацией свободных зарядов n. На фиг. 2 показана зависимость α от логарифма числа электронов n,



Фиг. 2

определенного по эффекту Холла. Прямая *1* вычислена по формуле (6). Эта зависимость оправдывается настолько хорошо, что отступление от нее можно рассматривать как указание на смешанный характер проводимости или на неоднородный состав образца. Так, точки кривой *2*

получены для различных соотношений смеси сернистого свинца с дырочной и электронной проводимостью, а точки кривой *3* — для Mg_3Sb_2, в котором запретная зона настолько узка, что в проводимости участвует сравнимое количество дырок и электронов.

Зависимость т. э. д. с. от концентрации электронов или от постоянной Холла имеет иное значение для вырожденных электронов металла. Действительно, для PbS вблизи $n = 10^{20}$ замечается излом прямой $\alpha = f(\log n)$, соответствующий вырождению.

В ранее изученной Жузе и Курчатовым закиси меди была получена зависимость электропроводности от температуры в виде

$$\sigma = A_1 e^{-\frac{\varepsilon_0}{2kT}} + A_2 e^{-\frac{\varepsilon}{2kT}},$$

где $\varepsilon_0 = 1.44$ эл.-в. и $\varepsilon \approx 0.6$ эл.-в.

Для сернистого свинца получается аналогичная зависимость с $\varepsilon_0 \approx 0.6$ эл.-в. и ε близким к нулю. Однако в этом случае ход проводимости σ не может характеризовать изменения числа носителей тока, так как $\sigma = neu$, а зависимость n и u от температуры одного порядка.

Измерения эффекта Холла позволили более строго разрешить эту задачу. При достаточно низких температурах можно пренебречь числом носителей тока, создаваемым решеткой PbS, по сравнению с теми, которые вызываются примесью свинца или серы. Оказалось, что n остается строго постоянным в широком интервале температур. Измеряя эффэкт Холла, Дунаев получил для одного образца: при 90° К $n = 1.0 \cdot 10^{9}$, при 573° К $n = 1.0 \cdot 10^{19}$.

Из этого совпадения можно заключить, что добавочные уровни энергии, вносимые свинцом или серой, лежат не в запрещенной зоне, а внутри свободной или заполненной полосы уровней.

Следует в таком случае ожидать, что число n не изменится и при отсутствии теплового движения и что примесная проводимость PbS сохранится вблизи абсолютного нуля. Действительно, измерения σ вплоть до температуры 2° К дали значительную проводимость порядка нескольких обратных омов. Проводимость убывала с повышением температуры, вследствие уменьшения подвижности.

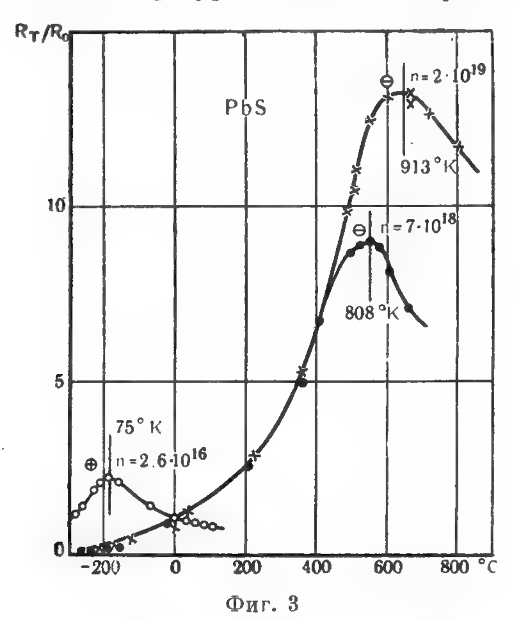
Утверждения о сверхпроводимости PbS не подтвердились ни при каком содержании растворенного свинца, когда он находился в «зака-

ленном» виде. Возможно, что бывают случаи, когда, коагулируя, примесь свинца образует сверхпроводящие мостики.

Проводимость при низких температурах целиком определяется концентрацией примеси свинца или серы.

На фиг. 3 изображен ход сопротивления с температурой для трех образцов сернистого свинца.

В первом случае концентрация примеси так мала, что еще значительно ниже температуры жидкого воздуха число электронов, освобождаемых решеткой, начинает превышать число электронов примеси. При 75° K рост числа электронов с температурой перевешивает падение подвижности, и проводимость PbS начинает расти.



Фиг. 3

Во втором образце, более богатом примесью, максимум сопротивления достигается только при 808° K, а в третьем с еще большим содержанием свинца при 913° K.

В первом образце концентрация добавочных электронов составляла $n=2.6\cdot10^{16}$, во втором $n=7\cdot10^{18}$ и в третьем $n=2\cdot10^{19}$. Можно получать образцы как с вырожденными электронами, так и с электронами, подчиняющимися статистике Максвелл — Больцмана. Если все растворенные атомы свинца или серы отдают или связывают по два электрона, то половина приведенных чисел дает число примесных атомов.

В тех пределах температур, где число электронов остается неизменным, кривые фиг. 3 дают точную зависимость подвижности электронов или дырок от температуры при различной их концентрации. При низких температурах этот ход достаточно хорошо может быть выражен формулой

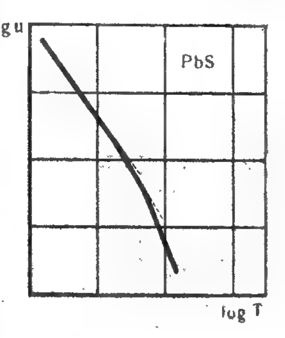
$$u=\frac{A}{T^{3/2}}.$$

С повышением температуры показатель степени при T возрастает. Нанося $\log u$ (или, что то же, $\log \sigma$) как функцию $\log T$ (фиг. 4), мы получаем для интервала температур от 2° K до комнатной $u \approx T^{-3/2}$, вблизи 100° C $u \approx T^{-2}$, при 200° C $u \approx T^{-2.7}$, а около 500° C $u \approx T^{-3.4}$. Так как концентрация электронов не превышает $2 \cdot 10^{19}$, то мы имеем дело с невырожденными электронами тепловых скоростей.

Следующая таблица дает концентрацию n электронов и величину u подвижности (в см/сек на 1 в/см) в пределах температур от 90 до 873° K.

T, °K	$n \times 10^{19}$	u	T, °K	$n \times 10^{19}$	u
90	1.0	2400	773	1.3	28.6
290	1.0	300	823	1.45	22.6
573	1.0	77	873	2.07	18.3
723	1.05	35.3			

При высоких температурах, далеких от минимума σ, рост проводимости определяется увеличением числа электронов. Из этой части кривой, введя поправку на подвижность, можно определить ширину запретной зоны ε_0 как в том случае, когда мы имеем дело с нормальными электронами, так и в случае вырожденных электронов.



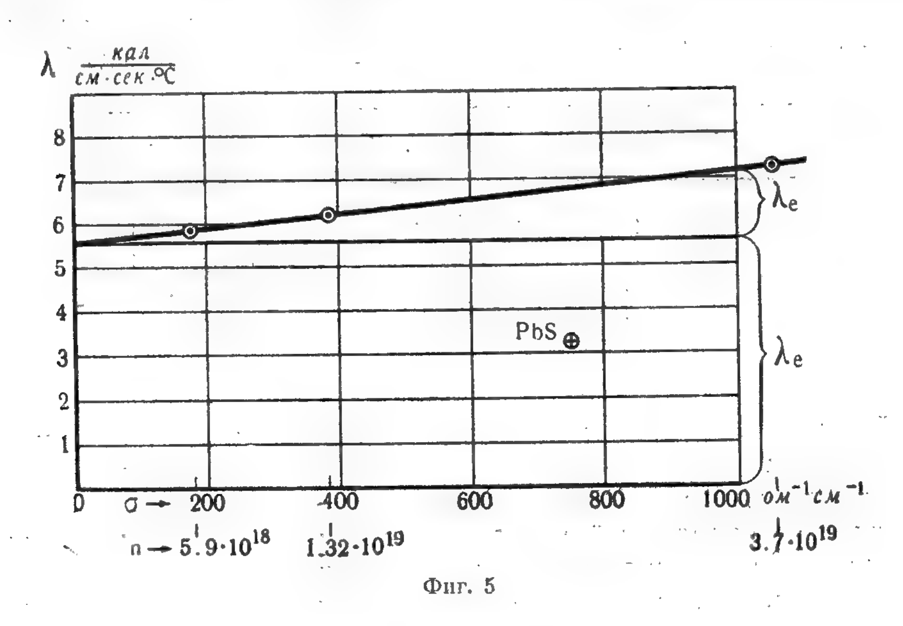
Фиг. 4

Таким образом, одна и та же кривая $\sigma = f(T)$ дает при низких температурах температурный ход подвижности, а при высоких — рост числа электронов.

Отметим еще один интересный результат, полученный Дунаевым и Маслаковцем для теплопроводности PbS. Ее зависимость от содержания примеси изображена на фиг. 5.

Горизонтальную часть можно считать равной теплопроводности решетки, а подъем в правой части приписать свободным электронам, создающим проводимость.

Добавочная теплопроводность $\Delta\varkappa$ пропорциональна электропроводности σ. Их отношение, как оказалось, удовлетворяет закону Видемана — Франца. Более точное измерение этого отношения должно установить различие между классическим его значением и величиной, справедливой для случая вырожденных электронов, а также установить значение $\Delta\varkappa/\sigma$ отдельно для дырок и электронов.

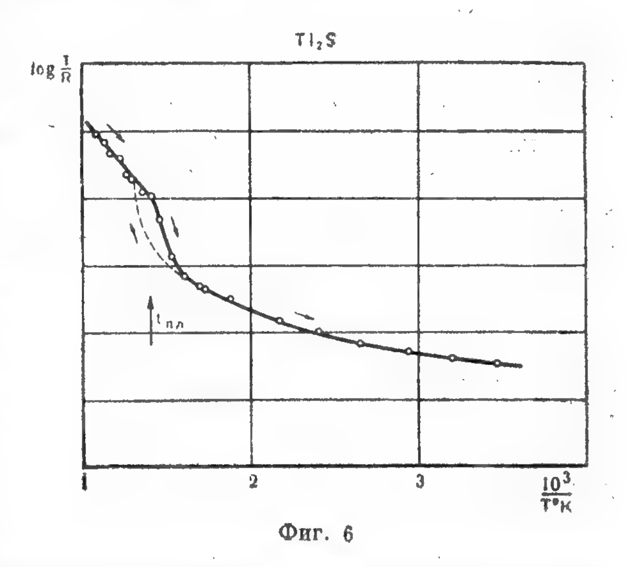


Фиг. 5

Итак, сернистый свинец представляет собою типичный полупроводник с примесью таких источников электронов или дырок, энергия которых лежит внутри полосы уровней основного вещества. Эти примесные электроны образуют металл с любой концентрацией электронов или дырок. В отличие от обычных металлов здесь можно иметь в свободной полосе любое число электронов, весьма далекое от вырождения. Эти электроны образуют металл классической физической теории Друде — Лорентца. Но можно, увеличивая концентрацию примеси, осуществить и исследовать также промежуточную область перехода к вырождению при еще большей концентрации примеси; наконец, получится вырожденный электронный газ в решетке диэлектрика.

Таким образом, изменением количества избыточного свинца или серы в сернистом свинце удается перекинуть мост между вырожденными электронами металла и больцмановской статистикой электронов обычных полупроводников.

Расплавленные полупроводники. Современная теория полупроводников связывает их свойства с кристаллической структурой. Проводимость приписывается полосе уровней кристаллической решетки. Каковы свойства аморфного или жидкого электронного полупроводника? Для них не существует ни теории, ни экспериментальных данных (за исключением недавно опубликованных опытов с жидким и аморфным селеном). А. Р. Регель измерил ход электропроводности ряда полупроводников (Tl_2S, Tl_2Se и др.) при переходе через температуру плавления. На фиг. 6 показаны полученные им результаты для Tl_2S. Пока нет достаточно надежных данных по эффекту Холла, толкование этих данных неполно.



Фиг. 6

Но можно утверждать, что электронная проводимость сохраняется в жидком состоянии, причем и после плавления продолжается рост концентрации носителей тока с температурой. Наклон прямых $\log \sigma = f\left(\frac{1}{T}\right)$, определяющих энергию диссоциации ε, не изменяется с переходом в расплав, хотя в жидкости трудно говорить о запрещенной зоне. Повидимому, величину ε разумнее называть работой ионизации или диссоциации электронов.

Подвижность уменьшается в момент плавления и в жидком состоянии очень мало меняется с температурой.

Термоэлектродвижущие силы в жидких полупроводниках имеют такой же порядок величины и тот же знак, что и в твердом состоянии, следовательно, механизм проводимости сохраняет тот же тип дырочного или электронного тока.

Распространение понятия о бриллуэновских зонах кристалла на аморфное тело вряд ли уместно. Система электронных уровней в жидкости может зависеть только от среднего расстояния между атомами флуктуирующего в значительных пределах. Электрические поля соседних атомов и вызываемые ими расщепления электронных уровней в различных атомах различны, резонанс сильно размыт, а электроны правильнее поэтому считать связанными с индивидуальными атомами. Тепловое движение диссоциирует часть связанных электронов, переводя их в свободное состояние. Наличие дырочной проводимости в жидких полупроводниках показывает, что и она не связана со строго периодической кристаллической структурой. Полная аналогия между свойствами твердых и жидких электронных полупроводников требует такого рассмотрения, которое охватывало бы оба состояния. Заметим, что и в твердых полупроводниках представление о диссоциации электронов с отдельных атомов часто лучше выражает физическое явление, чем строгая квантовая теория периодической решетки, объединяющая всю совокупность электронов.

Одна из этих картин хорошо описывает предельный случай металла, другая — изолятор. Для полупроводников обе картины отражают различные стороны наблюдаемых явлений.

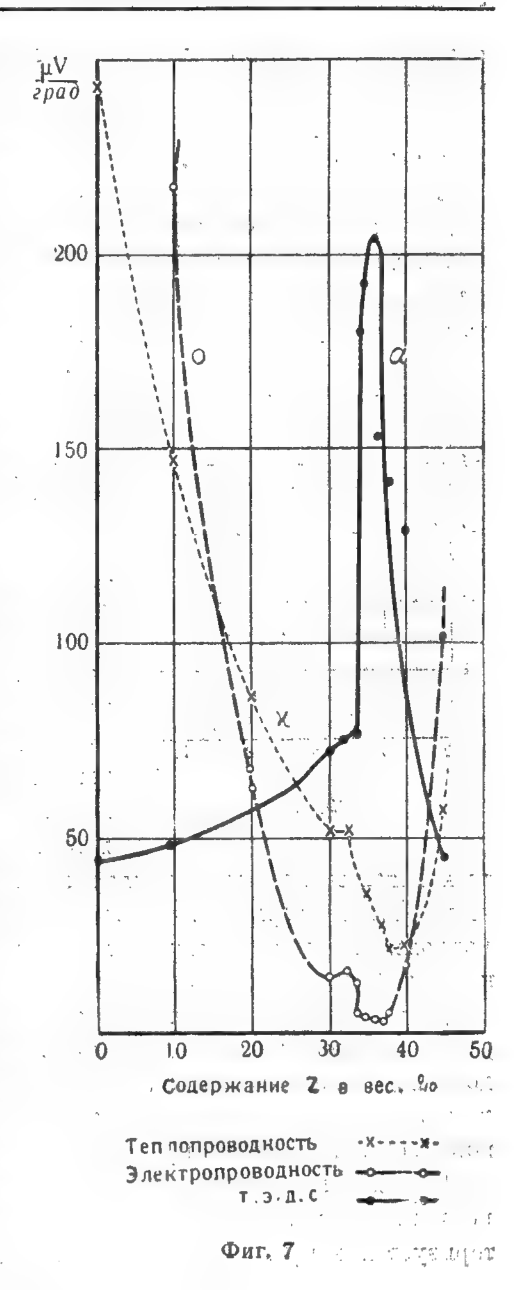
Сплавы металлов. Волновая механика проводит резкую грань между металлами и другими проводниками электричества. Металлы обладают незаполненной полосой энергетических уровней, в то время как во всех других материалах энергетические уровни валентных электронов отделены от полосы свободных уровней запретной зоной конечной ширины. Поэтому валентные электроны металла свободно проводят ток независимо от теплового движения, в том числе и вблизи абсолютного нуля температур. В электронных же и ионных полупроводниках носители тока освобождаются за счет тепловой энергии; поэтому их электропроводность стремится к нулю по мере приближения к абсолютному нулю температур.

Другое существенное отличие металлов от полупроводников вытекает из большой концентрации свободных электронов в металле.

Все уровни кинетической энергии от нуля до 5—10 эл.-в. оказываются занятыми, так что в переносе тока участвуют электроны высокой энергии. В полупроводниках тепловое движение переводит в свободную полосу сравнительно небольшое число электронов, энергия которых распределяется статистически так же, как среди молекул газа. При комнатной температуре средняя кинетическая энергия электронов полупроводника составляет всего 0.025—0.03 эл.-в.

Силы сцепления в металле вызываются взаимным притяжением всей совокупности свободных электронов с кристаллической решеткой положительных ионов. Металл принимает форму такой решетки, при которой энергия его электронов оказывается наименьшей.

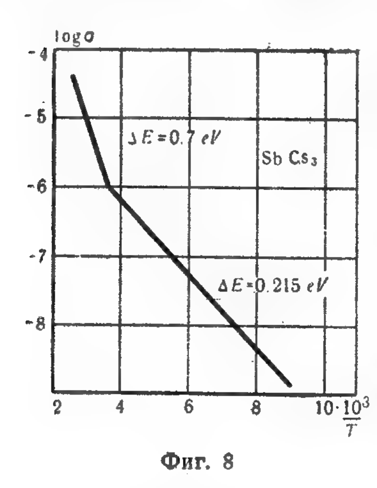
Это же требование определяет и свойства металлических сплавов с той разницей, что кристаллическая решетка сплавов состоит из нескольких видов положительных ионов, расположенных либо в строго периодическом порядке в упорядоченных сплавах, либо же в хаотическом беспорядке в неупорядоченных сплавах. Химическая природа ионов, из которых состоит решетка металла, отступает на задний план. Свойства



Фиг. 7

сплава определяются не валентностью атомов сплава, а соотношением между числом электронов и числом атомов в элементарной ячейке кристалла. Например, для всех β-латуней это отношение составляет 3/2, для γ-латуней 21/13, для гексагональных структур 7/4.

Примеси и отступления от стехиометрического состава служат источниками зарядов, переносящих ток, и поэтому резко повышают проводимость электронных и ионных полупроводников. В металлах, наоборот, влияние примеси сказывается лишь в уменьшении подвижности электронов благодаря рассеянию на них электронных волн.



Фиг. 8

Металлический блеск, сильное поглощение света, возрастающее с переходом в инфракрасную область спектра — естественное следствие наличия в металлах большой концентрации свободных электронов.

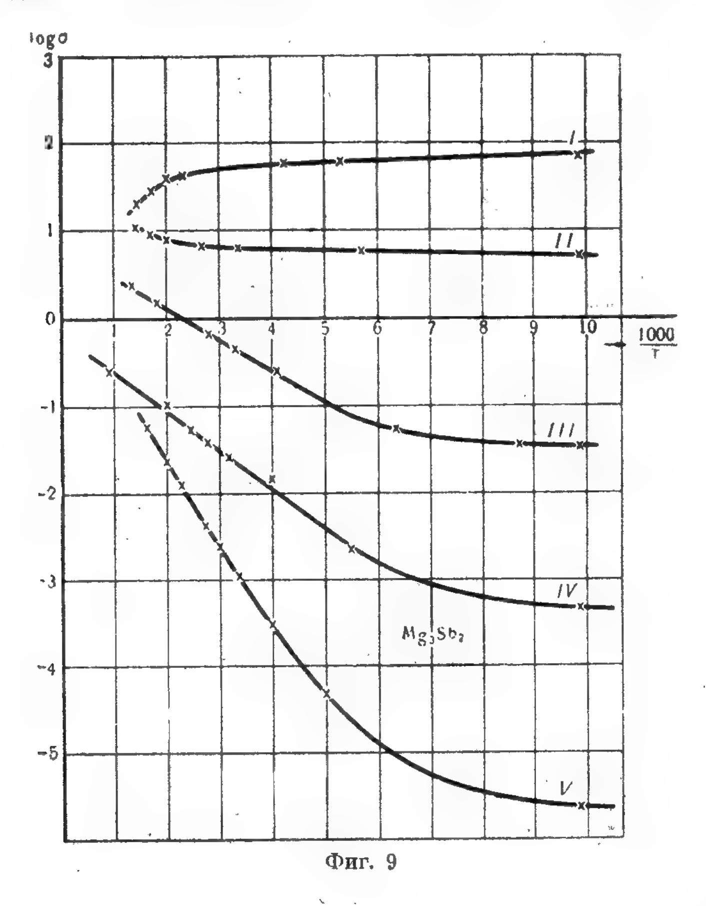
Таким образом, целый ряд признаков выделяет металлы и их сплавы из многообразия твердых и жидких материалов.

Н. С. Курнаков и его школа, всесторонне изучившие большое число металлических сплавов, обратили внимание на некоторые из них — «дальтониды», которые удовлетворяют стехиометрическим соотношениям и по своим свойствам напоминают обычные химические соединения. На диаграмме, выражающей зависимость температуры плавления от состава, дальтониды выделяются острыми максимумами. Более плавные максимумы, состав которых не может быть выражен правилами валентности, Курнаков назвал «бертоллидами».

Дальтониды и в меньшей степени бертоллиды отличаются высоким удельным сопротивлением, большой термоэлектродвижущей силой, твердостью, тугоплавкостью. Вопрос о том, следует ли их относить к металлам, не может считаться решенным, так как все эти отличия носят чисто количественный, а не принципиальный характер.

Решающим признаком металла, как мы видели, следует считать незаполненность полосы уровней валентных электронов, а диэлектрик характеризуется запретной зоной в системе энергетических уровней. Рассмот-

ренный нами ранее пример сернистого свинца показывает, что не всегда легко установить существование заполненной зоны. Примесь свинца или серы давала там свободные заряды, число которых не зависело

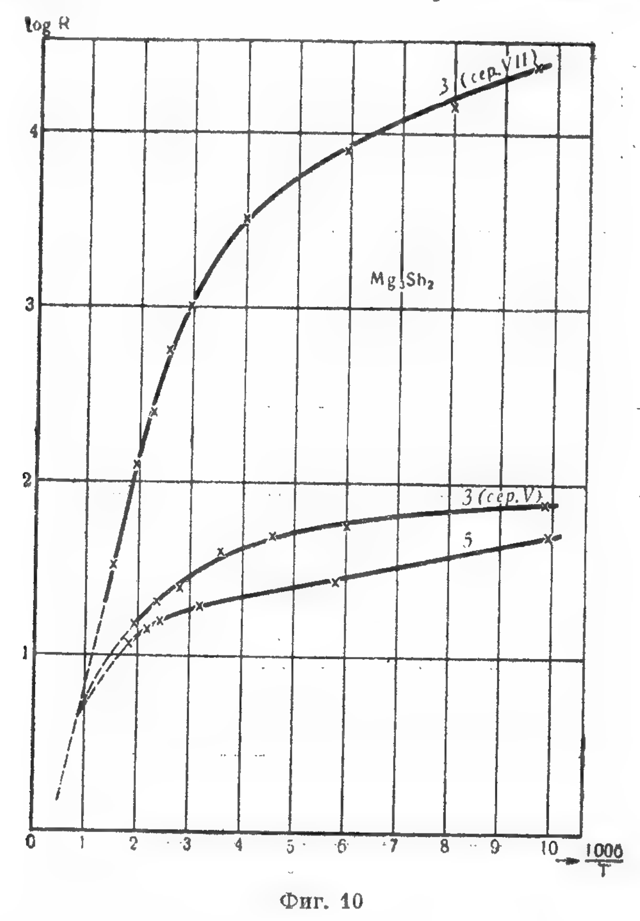


Фиг. 9

от температуры. Достаточно избытка свинца в 0.01%, чтобы получить больше электронов, чем может дать тепловое движение основной решетки PbS.

С повышением температуры число последних быстро возрастает, тогда как число электронов, создаваемых примесью, не изменяется. При высоких температурах соотношение резко меняется: число электро-

нов, создаваемых примесью, ничтожно по сравнению с тепловыми электронами. Кривые на фиг. 1 показывают, что в этой области температур $\log \sigma$ линейно убывает с $1/T$, а наклон прямой дает величину $\varepsilon_0/2$.



Фиг. 10

Поскольку при изготовлении сплава мы не можем совершенно точно обеспечить стехиометрический состав во всех участках образца, можно ожидать, что проводимость будет зависеть от небольших отступлений от точных соотношений компонентов. Однако, измерив проводимость σ, постоянную эффекта Холла R и термоэлектродвижущую силу α в ши-

роком интервале температур, мы можем отделить электроны примеси от электронов основной решетки сплава.

В. П. Жузе, Е. Д. Девяткова, И. В. Мочан, Б. И. Болтакс произвели такие измерения на сплавах Cs_3Sb, Mg_3Sb_2, $ZnSb$, Mg_2Sn, Mg_2Pb и некоторых других.

Кривые фиг. 7 дают зависимость σ, α и теплопроводности $\varkappa$ от состава для сплава $Zn = Sb$ по данным Е. Д. Девятковой. При соотношении компонентов, удовлетворяющих указанным формулам, получается резкий минимум удельной проводимости, достигающий в случае Mg_3Sb_2 10^{-4} ом$^{-1}$ см$^{-1}$, т. е. значения в 10^9 раз меньше, чем в чистых металлах.

Зависимость $\log \sigma$ и $\log R$ от $1/T$ дана на кривых фиг. 8, 9, 10. Из них можно заключить, что во всех перечисленных сплавах существует запретная зона уровней, ширина которой, судя по температурному ходу постоянной эффекта Холла и электропроводности, равна:

для Cs_3Sb	ε_0 около 0.7 эл.-в.	ε для примеси 0.2 эл.-в.
Mg_3Sb_2	0.5	0
$ZnSb$	0.4	0.07
Mg_2Sn	0.2	0

Экстраполируя ход $\log \sigma = f\left(\frac{1}{T}\right)$ к комнатным температурам, можно утверждать, что при точном стехиометрическом составе исследованные нами сплавы должны обладать электропроводностью: Cs_3Sb — ок. 10^{-7} ом$^{-1}$ см, Mg_3Sb_2 — ок. 10^{-4}, $ZnSb$ — ок. 10^{-3}, Mg_2Sn — ок. 10^{-1} ом$^{-1}$ см.

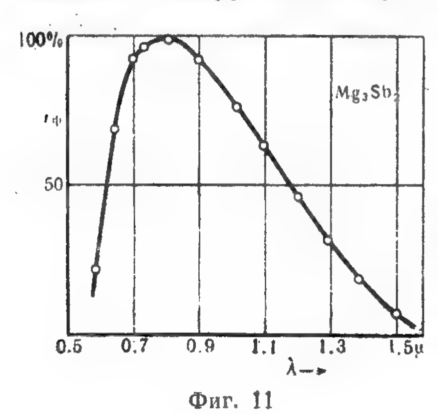
Эта электропроводность и соответственная концентрация электронов уменьшаются с понижением температуры и падают до нуля с приближением к абсолютному нулю температур.

Таким образом, не может быть и речи о том, чтобы такое ничтожное число электронов могло связывать сплав и давать ему свойства твердого тела. Несомненно, что силы сцепления в этих сплавах имеют иное происхождение, совпадающее с валентными связями изоляторов.

Убедительным и наглядным проявлением запретной зоны служит явление фотопроводимости, при котором преодоление энергетического

барьера совершается за счет энергии поглощенного фотона. С. М. Рывкин и И. В. Мочан действительно установили присутствие фотопроводимости в ZnSb и Mg_3Sb_2 (фиг. 11). Еще отчетливей фотопроводимость в Cs_3Sb — веществе наиболее чувствительных фотокатодов.

Наилучший метод изучения сплавов и зависимости различных их свойств от состава предложен был С. А. Векшинским. Два металла возгоняются в вакууме на одну и ту же пластинку из нагревателей, расположенных на противоположных ее концах. Состав осадка непрерывно меняется от одного конца пластинки к другому.



Фиг. 11

На каком-то участке его состав близко подходит к стехиометрическому. Такой методикой воспользовалась И. В. Мочан. Можно было видеть, что на соответственном участке сплав лишается металлического блеска, становится сравнительно прозрачным, в особенности для инфракрасных лучей. Здесь наблюдается фотопроводимость и повидимому, также перенос вещества током (электролиз).

В полном соответствии с уменьшением концентрации свободных электронов исчезает не только электропроводность, но и вызванная ими большая теплопроводность металлов. Так, с приближением к стехиометрическому сплаву Mg_3Sb_2 и ZnSb удельная теплопроводность падает от нескольких десятых до нескольких тысячных.

Отличие сплавов, названных Н. С. Курнаковым дальтонидами, от металлов настолько глубоко, что сплавы эти необходимо отнести к валентным химическим соединениям. В случаях Mg_3Sb_2 и Cs_3Sb очень вероятно, что валентные электроны Mg и Cs переходят к Sb, образуя полярную решетку из положительных ионов Mg^{++} или Cs^{+} и отрицательных ионов Sb^{---}. В сплавах металлов, более близких по своему ионизационному потенциалу, возможна валентная связь парами электронов с противоположными спинами. Вообще же, среди сплавов можно ожидать появления всех тех видов кристаллических решеток, которые известны для диэлектриков.

Но можно пойти и дальше. Можно предположить, что и в истинных металлах с незаполненной полосой уровней наряду с силами сцепления, вызванными притяжением между решеткой положительных ионов и свободно движущимися между ними электронами, действуют и химические связи между атомами. Действительно, энергия электронов уменьшается вблизи положительных ионов; поэтому при своем хаотическом движении они больше времени проводят вблизи ионов, чем между ними. В эти периоды времени электрон составляет часть атома, занимая его квантовую орбиту (хотя и искаженную полем других атомов), и связывает его с соседними атомами обычными валентными связями.

Спектры поглощения и испускания мягких рентгеновых лучей показали, что электронам в металлах необходимо приписать не только определенную квантовую кинетическую энергию, но и остальные квантовые числа, свойственные электрону в атоме или молекуле газа. Предложенная еще в 1926 г. Я. И. Френкелем теория металлов (которая впервые приписала электронам в металле квантовую энергию) рассматривала движение электронов как статистический переход с квантовой орбиты одного атома на соседнюю такую же орбиту. Такое представление об электронах в металле столь же закономерно, как и гипотеза Зоммерфельда — Бете о свободном движении электронных волн сквозь периодическую кристаллическую решетку. Последнее представление описывает металлическую связь, теория же Френкеля делает понятной также проявление валентных сил связи в металле.

Делингер объясняет кристаллизацию щелочных металлов Na, K в решетку центрированного куба вместо наиболее плотной решетки с центрированными гранями обменной связью двух соседних атомов. Присутствие значительного количества (свыше 1%) двухатомных молекул в парах натрия подкрепляет предположение, что и в металле в какой-то степени проявляется тенденция к объединению соседних пар электронов с противоположными спинами.

Таким образом, мы приходим к представлению, что даже в чистых металлах и в истинно металлических сплавах, наряду с металлическими связями между совокупностью ионов и электронов, хотя бы временно проявляются и чисто химические взаимодействия между атомами.

Дальтониды и бертоллиды представляют с этой точки зрения лишь предельные случаи чисто валентных связей. В общем случае мы имеем

сочетание металлических и валентных связей в самых различных соотношениях.

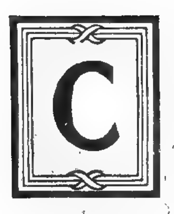
Выводы. Теория полупроводников развивалась под влиянием представлений, созданных квантовой теорией кристаллических металлов и изоляторов. Изучение полупроводников в расплавленном состоянии и промежуточных типов хорошо проводящих полупроводников и плохо проводящих сплавов значительно расширяет область электрических свойств вещества. Оказывается, что и классические представления о диссоциации и рекомбинации электронов, выросшие из наблюдений над электрическим поведением газов, сохраняют свое значение также для аморфных и кристаллических полупроводников.

Установлен критерий для суждения о том, какие вещества следует относить к металлам. Многие сплавы металлов с этой точки зрения оказываются примесными полупроводниками с ионной или молекулярной решеткой. Ряд соображений и вновь установленных фактов приводит к гипотезе, что и в истинных металлах наряду с металлическими силами сцепления в большей или меньшей степени проявляются и валентные химические силы.

Академик

Г. С. Ландсберг

МОЛЕКУЛЯРНОЕ РАССЕЯНИЕ СВЕТА



Среди вопросов молекулярной физики, успешно разрешенных за последние годы, вопросы, касающиеся строения молекул, занимают одно из самых первых мест. Проблемы этого рода одинаково важны как для физики, так и для теоретической химии и даже для химии аналитической. Современная физика нашла разнообразные методы, позволяющие судить о геометрических параметрах, характеризующих пространственное распределение атомов (и ионов), составляющих молекулу, и о параметрах динамических, определяющих величину и характер внутримолекулярных сил.

Приемы исследования размеров молекул и взаимного расположения атомов в молекулах и кристаллических структурах развивались уже давно, но совершенно особую роль сыграл в этом отношении метод, основанный на диффракции X-лучей, давший возможность точного и уверенного определения величин, характеризующих *статику* молекулы: положение, взаимное расстояние и характер групп атомов или ионов, ее составляющих.

Применение этого метода быстро обогатило науку сведениями, относящимися к архитектуре материи, в первую очередь к архитектуре кристаллического состояния. Поэтому 1912 год, дату открытия диффракции X-лучей, следует по праву считать решительным этапом в этой области. Позже к этому методу присоединился метод, основанный на диффракции

электронов, приведший в настоящее время к созданию электронного микроскопа.

Совершенно аналогично обстоит дело и с вопросами молекулярной динамики. Различные методы оценки внутримолекулярных сил, в частности исследования инфракрасного поглощения, были известны давно, но быстрое накопление данных в этой области и их широкая и плодотворная интерпретация начались лишь с 1928 г., когда было установлено явление комбинационного рассеяния света.

Для понимания сущности явления комбинационного рассеяния, открытого в конце февраля 1928 г. одновременно в работах Л. И. Мандельштама и автора настоящего очерка и в работах индусских физиков Рамана и Кришнана, необходимо дать представление о молекулярном рассеянии света вообще.

Яркий пучок, например пучок от прожектора, можно видеть «со стороны». Это значит, что свет не только распространяется по направлению, заданному оптической системой («по оси прожектора»), но и рассеивается в стороны. Это явление наблюдается весьма отчетливо в запыленном воздухе и вообще в мутной среде. Вернее, само понятие «мутная среда» означает, что в ней наблюдается сильное рассеяние света. В простейшем случае мутность обусловлена тем, что среда представляет собою смесь разнородных частиц, оптические свойства которых различны, т. е. среда оптически неоднородна.

Оптическая неоднородность легко может быть получена в газах (пыль, туман, дым), в жидкостях (суспензии и эмульсии, например, разведенное молоко), в твердых телах (молочное стекло). Исследование мутных сред такого рода имеет практическое значение (ориентировка в тумане, нефелометрия, применение молочных стекол в светотехнике). С теоретической же стороны понять это явление не трудно и оно не представляет особого интереса, ибо физическая причина мутных свойств таких сред зависит от случайных явлений — образования искусственных неоднородностей.

Вопрос приобретает несравненно больший теоретический интерес, если мы переходим к исследованию сред, вполне свободных от этих загрязнений и представляющихся в смысле чистоты идеальными. В этом случае оптическая неоднородность, если она имеется, должна обусловливаться только строением вещества, так что изучение

рассеяния света в таких вполне чистых веществах есть метод исследования их строения.

В самом деле, первое теоретическое рассмотрение подобного вопроса, выполненное Релеем еще в 80-х годах прошлого столетия, связало явление рассеяния света в чистых газах с молекулярным строением их. Релей полагал, что взаимодействие световой волны с молекулами воздуха должно вести к рассеянию света, причем на основании его вычислений можно было сделать два важных заключения.

Во-первых, из теории Релея следует, что сильнее рассеивается свет более короткой длины волны (синий рассеивается сильнее красного), так что интенсивность рассеяния обратно пропорциональна четвертой степени длины волны $\left(\text{закон } \frac{1}{\lambda^4}\right)$.

Это заключение позволило объяснить голубой цвет неба, ибо свет небесного свода есть свет солнца, рассеянный воздухом атмосферы и, по закону Релея, значительно обогащенный синими и фиолетовыми лучами по сравнению с прямым светом солнца. Наоборот, солнце и луна на восходе и закате кажутся багрово-красными, потому что из их прямых лучей до нас доходят главным образом наиболее длинные волны (красные), так как в толстом слое атмосферы более короткие волны сильно рассеиваются.

Второй вывод Релея позволил связать интенсивность рассеянного света с числом молекул в единице объема. Этот вывод сделал возможным из наблюдений над интенсивностью света неба определить число молекул воздуха в одном кубическом сантиметре, т. е. найти чрезвычайно важную для молекулярных представлений постоянную.

Таким образом, уже на этом примере ясно теоретическое значение явления молекулярного рассеяния. С экспериментальной стороны исследование рассеяния света — трудная проблема, так что надежные результаты, подтвердившие расчеты Релея и позволившие определить число молекул в единице объема по интенсивности света неба, относятся лишь к 1906—1910 гг.

К этому же времени относится важное теоретическое изменение, внесенное в первоначальную теорию Релея. Л. И. Мандельштам указал, что для понимания явления рассеяния недостаточно обычных представлений беспорядочного молекулярного движения, которым поль-

зовался Релей. Причина оптической неоднородности не может лежать в самом факте существования молекул, ибо размеры молекул очень малы по сравнению с длиной световой волны, а для рассеяния света необходимо, чтобы размеры неоднородностей были сравнимы с длиной волны.

Для понимания явления необходимо было учесть, что молекулярное движение способно вызвать скоропреходящие скопления и разрежения молекул, образующие неоднородности (уплотнения), размер которых достаточен для рассеяния света. На роль этих скоплений и разрежений (флюктуации плотности) для исследуемого явления указал Смолуховский, сделавший, таким образом, следующий шаг вперед. Он применил свои соображения к объяснению одного давно известного случая сильного рассеяния, так называемой опалесценции при критической температуре, а Эйнштейн и Мандельштам развили на основе идеи Смолуховского теорию рассеяния света внутри вещества (Эйнштейн) и на поверхности его, на границе двух сред (Мандельштам).

Таким образом, теория молекулярного рассеяния света была поставлена на твердую почву. Вслед за этим начались усиленные попытки экспериментального изучения явления в лабораторной обстановке. Ряд исследователей (Кабанн — во Франции, Ганс — в Лаплате, затем в Германии, Раман — в Индии, Стретт — в Англии, Мартин — в Америке) получили вполне определенные результаты по исследованию рассеяния света в газах и жидкостях и выяснили чрезвычайно важные особенности явления, позволявшие проверить и углубить теорию.

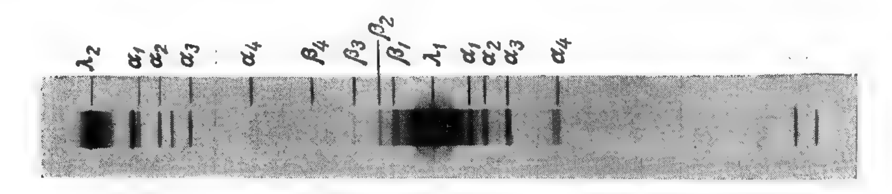
Однако во всех этих исследованиях, равно как и во всех теоретических построениях, с ними связанных, вопрос о спектральном составе рассеянного света не затрагивался по существу. Теория Релея говорила лишь об изменении распределения энергии в спектре рассеянного света по сравнению с первичным светом, указывая на относительное увеличение интенсивности более коротковолновой части спектра. Эксперимент подтверждал общий ход изменения этого спектрального распределения. Если представить себе, что на среду направляется монохроматический свет, т. е. свет определенной частоты, то первоначальная теория не позволяла ожидать никаких изменений: действительно, рассеянный свет есть свет, испускаемый молекулами среды, совершающими колебания под действием падающего на них света. Эти вынужденные колебания должны



Л. Мандельштам

совершаться с частотой, навязанной им внешней силой, т. е. рассеянный свет должен повторять частоту падающего света.

Однако спектральное исследование рассеянного света, предпринятое нами и индусскими исследователями, показало, что дело обстоит гораздо сложнее. В случае освещения вещества монохроматическим светом частоты ν, рассеянный свет обнаруживает наряду с этой первоначальной частотой ν также частоты $(\nu+m_1)$ и $(\nu-m_1)$, $(\nu+m_2)$ и $(\nu-m_2)$ и т. д., где m_1, m_2 — определенные, характерные для изучаемого вещества величины.



Фиг. 1

На фиг. 1 показан вид спектра света, рассеянного четыреххлористым углеродом. На снимке видно наличие четырех добавочных частот, соответствующих четырем значениям величин m_1, m_2, m_3 и m_4 (последняя частота состоит из двух близких частот, так что линия, ей соответствующая, имеет вид двойной линии).

Исследование нового явления на ряде объектов позволило установить следующие основные законы его:

1. Каждая монохроматическая линия падающего света вызывает в данном веществе две системы симметричных сателлитов, с частотами $(\nu-m_1)$ и $(\nu+m_1)$, $(\nu-m_2)$ и $(\nu+m_2)$, $(\nu-m_3)$ и $(\nu+m_3)$ и т. д., причем значения m_1, m_2, m_3, ... не зависят от частоты падающего света ν, но характерны для исследуемого вещества. Табл. 1 иллюстрирует эту закономерность. При рассмотрении этой таблицы, равно как и дальнейшего материала, надо иметь в виду следующее. Во всех теоретических формулах мы всегда имеем в виду **частоты**, т. е. число колебаний в **секунду** (сек$^{-1}$). Результаты же спектроскопических измерений выражаются обычно в волновых числах, показывающих, сколько волн укладывается на сантиметре (см$^{-1}$). Связь между этими величинами очень

проста: волновое число равно частоте, разделенной на скорость света ($c = 3 \cdot 10^{10}$ см/сек).

Таблица 1

ЗНАЧЕНИЯ ВОЛНОВЫХ ЧИСЕЛ ДЛЯ КРАСНЫХ И ФИОЛЕТОВЫХ САТЕЛЛИТОВ

Возбуждающая линия	Хлороформ $CHCl_3$		Четыреххлористый углерод CCl_4	
	красный сателлит	фиолетовый сателлит	красный сателлит	фиолетовый сателлит
33086	667.7	—	459.7	—
31985	667.7	669.3	458.3	459.4
31923	667.3	669.5	457.8	460.4
29917	668.1	670.5	458.7	458.8
27354	666.8	669.9	459.4	461.8
27290	—	669.8	459.1	461.8
24705	669.2	669.5	457.9	459.7
23034	667.3	—	—	—
22938	669.4	667.1	459.6	457.8
Среднее	668.0	669.4	458.5	460.2

2. Интенсивности сателлитов, соответствующих увеличенной частоте (расположенных ближе к фиолетовому концу спектра), значительно слабее, чем интенсивность соответственных «красных» сателлитов, причем это различие интенсивностей тем больше, чем больше соответствующее m. Фиг. 1 хорошо иллюстрирует эту закономерность, ибо в случае четыреххлористого углерода интенсивность красных сателлитов, соответствующих m_1, m_2 и m_3, случайным образом почти одинакова, и поэтому уменьшение интенсивности фиолетовых сателлитов по мере увеличения m видно очень отчетливо.

3. Значения m_1, m_2, m_3, ..., характеризующие сдвиг частоты линий комбинационного рассеяния по сравнению с частотой возбуждающей линии, представляют собою не что иное, как значение частот собственных колебаний молекул изучаемого вещества, известных по линиям поглощения инфракрасных спектров. Эта важная закономерность иллюстрируется табл. 2.

Таблица 2

СОПОСТАВЛЕНИЕ ВОЛНОВЫХ ЧИСЕЛ ПО ДАННЫМ ИНФРАКРАСНОГО ПОГЛОЩЕНИЯ И КОМБИНАЦИОННОГО РАССЕЯНИЯ

Толуол		Кристаллический кварц		Толуол		Кристаллический кварц	
инфра-красн.	комбинац.	инфра-красн.	комбинац.	инфра-красн.	комбинац.	инфра-красн.	комбинац.
—	210	128	127*	1190	1200	800	800
—	515	—	208*	1380	1375	1110**	1075
—	620	263**	268	1456**	—	1200**	1250
767	780	384	360	1550	—		
979	981**	—	405	1610	1600		
1030**	1024	480	465**	3000*	2920*		
1080**	—	—	694		3050*		

Примечание. Одной звездочкой отмечены сильные и двумя — очень сильные линии.

Ряд отступлений, бросающихся в глаза при рассмотрении этих таблиц и состоящих в том, что многим линиям комбинационного рассеяния не соответствуют линии инфракрасного поглощения, отчасти объясняется неполнотой инфракрасных спектров, отчасти имеет другие, более глубокие причины, которые будут рассмотрены ниже.

Отвлекаясь пока от этих осложнений, мы уже здесь должны подчеркнуть, что установленная закономерность дает возможность по спектрам комбинационного рассеяния света устанавливать частоты собственных колебаний молекул, причем этот новый метод имеет огромные преимущества перед существовавшими ранее методами исследования инфракрасных спектров поглощения молекул.

Преимущества эти заключаются в технической простоте нового метода определения по сравнению с методикой инфракрасного поглощения, особенно в области сравнительно длинных инфракрасных волн, где лежит большинство собственных частот молекул. Уже одно это преимущество сразу позволило значительно расширить круг соответствующих исследований и очень быстро обогатило науку обширным материалом, относящимся к собственным колебаниям молекул, открыв, таким образом,

22*

широкий доступ к изучению вопросов молекулярной динамики. Не менее важным преимуществом является также гораздо бо́льшая разрешающая способность нового метода, позволяющего обнаруживать ряд важных деталей, относящихся к собственным колебаниям молекул, ускользавших при методе инфракрасного поглощения. Очень важно также, что метод комбинационного рассеяния сразу дает о с н о в н ы е собственные колебания, тогда как по методу инфракрасного поглощения получаются одновременно и обертоны и вторичные (разностные и суммовые) частоты, что очень затрудняло интерпретацию спектров и нередко делало ее неоднозначной.

Впрочем, следует теперь же отметить, что метод комбинационного рассеяния вовсе не исключает метод инфракрасного поглощения. Оба метода успешно дополняют друг друга, так что для углубленного теоретического анализа нередко бывает весьма полезно наряду с данными исследований спектра комбинационного рассеяния привлекать также и данные по инфракрасному поглощению.

Перечисленные три основные закономерности могут быть записаны в виде простого равенства

$$\nu' = \nu \pm m_k, \tag{1}$$

где ν' — частота рассеянного света, ν — частота падающего света и m_k — частоты собственных колебаний молекул рассеивающего вещества. Так, например, для четыреххлористого углерода имеем: $m_1 = 6.5 \cdot 10^{12}$, $m_2 = 9.9 \cdot 10^{12}$, $m_3 = 13.8 \cdot 10^{12}$, $m_4' = 22.8 \cdot 10^{12}$ и $m_4'' = 23.7 \cdot 10^{12}$. В этих равенствах заключается одновременно и упрощенная наглядная теория явления, которую легче всего выразить на языке световых квантов.

Развитие идеи квантов привело к представлению, что всякий периодический процесс частоты ν характеризуется порциями энергии, величина которых равна $h\nu$, где h — универсальная постоянная, равная $6.55 \cdot 10^{-27}$ эрг·сек и носящая название постоянной Планка. Умножив равенство (1) на h, получим соотношение

$$h\nu' = h\nu \pm hm_k, \tag{2}$$

которое представляет не что иное, как общеизвестный закон сохранения энергии применительно к рассматриваемому явлению рассеяния света.

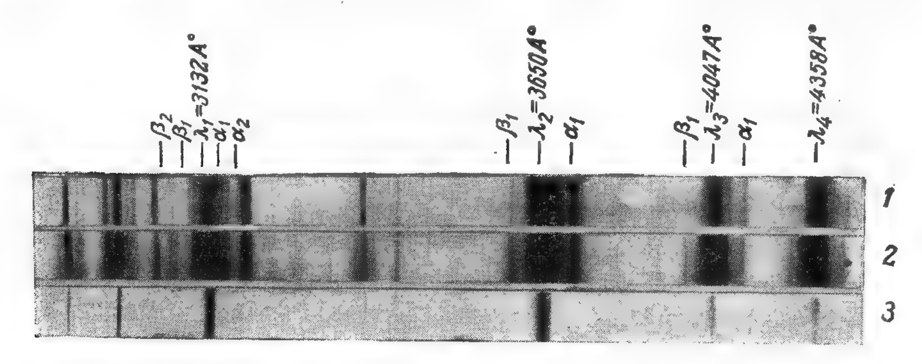
Действительно, интерпретация нашего явления в духе представления о световых квантах состоит в рассмотрении явления рассеяния как процесса соударения квантов света (с энергией $h\nu$) с молекулами рассеивающего вещества, характеризующимися частотой m_k и, следовательно, могущими захватывать или испускать квант энергии hm_k.

Если в процессе такого соударения молекула захватит порцию hm_k, то световой квант отразится (рассеется) с уменьшенной энергией $h\nu' = h\nu - hm_k$, т. е. с уменьшенной частотой (образование красного сателлита). Если же молекула при соударении отдает имеющийся в ней квант hm_k, то рассеянный квант будет обогащен энергией, и рассеяние произойдет согласно отношению $h\nu' = h\nu + hm_k$, т. е. образуется фиолетовый сателлит с увеличенной частотой $\nu' = \nu + m_k$.

Эта простая интерпретация позволяет в то же время легко понять, почему интенсивность фиолетовых сателлитов гораздо меньше, чем красных, и это отношение быстро уменьшается при увеличении m_k. В самом деле, с рассматриваемой точки зрения интенсивность сателлита рассеянного света определяется числом актов взаимодействия между квантом света и рассеивающей молекулой, причем красный сателлит может быть получен при взаимодействии света с любой невозбужденной (не колеблющейся) молекулой, тогда как для получения фиолетового сателлита необходимо взаимодействие света с молекулой возбужденной, обладающей добавочной энергией hm_k. Отношение числа возбужденных N' и невозбужденных N молекул определяется законом Больцмана, который в основном сводится к соотношению $\frac{N'}{N} = e^{-\frac{hm}{kT}}$, где $k = 1{,}36 \cdot 10^{-16}$ эрг/град — так называемая постоянная Больцмана, а T — абсолютная температура среды. Вводя значения h, k, T и m для условий, имеющих место на опыте, убедимся, что N'/N значительно меньше единицы, особенно для больших m, т. е. для сателлитов, далеко отстоящих от первичной линии. Таким образом, условия для образования фиолетовых сателлитов должны встречаться значительно реже, чем для образования красных, особенно для удаленных сателлитов. Красные сателлиты должны быть гораздо интенсивнее, чем соответствующие фиолетовые, и различие в интенсивностях тем больше, чем дальше расположены сателлиты от основной линии.

Соотношение $\frac{I_ф}{I_k} = \frac{N'}{N} = e^{-\frac{h\nu}{kT}}$ указывает в то же время, что отношение интенсивности фиолетовых сателлитов к красным должно быстро увеличиваться при возрастании температуры T рассеивающего вещества.

Опыты с кристаллами, допускающими значительное повышение температуры, в этом отношении могут быть особенно показательными. На фиг. 2 воспроизведен результат одного из опытов по рассеянию света в кварце, поставленных нами с целью исследования влияния



Фиг. 2

температуры на интенсивность линий комбинационного рассеяния. Он показывает, как значительно возрастает относительная интенсивность фиолетовых сателлитов кварца при увеличении температуры от 290 до 500° К.

Измерения интенсивности фиолетовых и красных сателлитов, равно как и основных линий рассеяния, выполненные нами, показали согласие с полной теорией, позволяющей предвидеть температурный ход интенсивности для всех этих линий, развитой в работах Л. И. Мандельштама, Г. С. Ландсберга, М. А. Леонтовича и И. Е. Тамма.

Наглядная интерпретация с точки зрения представления о световых квантах передает наиболее яркие черты явления, однако она не углубляется в механизм рассеяния света, ограничиваясь формальными заключениями. В соответствии с этим остаются неосвещенными другие, не менее важные стороны явления. Например, совершенно не затрагивается вопрос, почему некоторые частоты, обнаруживаемые по методу инфра-

красных колебаний, не проявляются в спектре комбинационного рассеяния (ср. табл. 2), и наоборот. Понимание этой стороны дела требует углубленной теории явления, теории, с помощью которой можно дать себе отчет об особенностях молекулярной динамики, и в соответствии с этим суметь по данным эксперимента сделать далеко идущие заключения о строении молекул. Для этой цели формализм теории световых квантов явно недостаточен, и задача потребовала бы углубленной квантово-теоретической трактовки. Однако такая трактовка мало подходит для создания наглядных представлений о существе явления; к тому же подобной квантовой интерпретации явления комбинационного рассеяния света во всей его полноте в настоящее время еще не существует.

С другой стороны, важнейшие моменты интересующего нас явления могут быть выяснены, если пользоваться чисто классическими представлениями. При этом, конечно, нельзя получить правильных количественных соотношений, ибо явление взаимодействия света и вещества носит по существу квантовый характер и для уточнения результатов необходимо внесение квантовых поправок. Несмотря на такой неудовлетворительный с теоретической точки зрения дуализм, подобный подход к явлениям комбинационного рассеяния света может быть оправдан. Он особенно пригоден для наших целей, ибо позволяет связать явление комбинационного рассеяния с явлением молекулярного рассеяния вообще и выясняет генезис представлений, приведших Л. И. Мандельштама и меня к открытию описываемого явления.

С классической точки зрения воздействие световой волны на вещество сводится к следующему. Под действием электрического поля световой волны атомы или молекулы вещества поляризуются, т. е. положительные и отрицательные заряды, их составляющие, сдвигаются в разные стороны. Таким образом, атом (молекула) приобретает электрический момент, который изменяется периодически в темпе изменения электрического поля волны.

Периодическое изменение электрического момента атомов (молекул) является источником вторичных электромагнитных волн, и взаимодействие этих волн между собою и с волной падающей определяет всю дальнейшую судьбу света: отражение, преломление или рассеяние.

Таким образом, все интересующие нас явления покоятся на колебательных процессах, на их возбуждении, их взаимодействии с другими

колебаниями или с волнами, ими вызванными. Поэтому мы должны рассмотреть некоторые типичные особенности волновых явлений.

Колеблющиеся системы обнаруживают особенно сильное взаимодействие в тех случаях, когда периоды их колебаний совпадают, т. е. когда они находятся в резонансе. Явление резонанса особенно чувствительно к расстройке между периодами колеблющихся систем, ибо такая расстройка аннулирует все преимущества воздействия в такт. В случае резонанса каждый следующий толчок, попадая в такт, помогает предыдущему; при расстройке же воздействия первого резонатора время от времени оказываются в оппозиции с движением второго и аннулируют, таким образом, эффект, произведенный ранее. Не трудно заметить, что если разность частот взаимодействующих систем равняется m в секунду, то усиление и ослабление воздействия также будут наступать m раз в секунду. Но если мы каким-либо образом устраним последовательно все невыгодные воздействия, т. е. в нашем случае m раз в секунду будем препятствовать первому осциллятору (например, камертону) оказывать тормозящее действие на движение второго, то возбуждающее действие первого проявится в полной мере, несмотря на расстройку в периодах.

Камертон, делающий ν колебаний в секунду и прерываемый m раз в секунду, входит в резонанс с другим камертоном, собственная частота которого отличается на m, т. е. равна $(\nu + m)$ или $(\nu - m)$.

Это важное заключение может быть выведено более строгим и общим способом при помощи математического анализа явлений, и оно вполне подтверждается опытом.

Воздействие на колеблющуюся систему, состоящее в периодическом прерывании ее колебаний или периодическом изменении интенсивности их, носит название модуляции. Мы можем, следовательно, сказать, что модулированное колебание эквивалентно существованию нескольких колебаний измененной частоты.

Действительно, пусть явления, имеющие место в генерирующей системе, неизвестны, и мы устанавливаем частоту этой системы путем настройки приемной аппаратуры, отвечающей на сигналы, т. е. настройки в резонанс (обычная радиотехническая практика применения «волномера» для определения частоты приходящих радиоволн). В таком случае мы можем лишь обнаружить, что полученные сигналы соответствуют частоте $\nu' = (\nu + m)$ и $\nu'' = (\nu - m)$; но мы не в состоянии, вообще го-

воря, отличить, являются ли ν' и ν'' собственными колебаниями двух генерирующих систем или мы имеем дело с модулированным колебанием какой-то одной системы, имеющей частоту, равную $\frac{\nu'+\nu''}{2}$, и подвергающейся периодическим воздействиям.

Не трудно обнаружить подобную модуляцию акустических волн с помощью двух расстроенных камертонов. Достаточно ослаблять или устранять в подходящем темпе волну, посылаемую одним из камертонов, закрывая при помощи регулярного движения руки отверстие резонансного ящика, на котором обычно укрепляется камертон; тогда приемный камертон (резонатор) обнаружит присутствие модулированных колебаний, отвечая на колебания измененной частоты.

Аналогичные эффекты изменения частоты в оптических опытах, т. е. изменение спектра света под действием модуляции световых колебаний, было бы, наоборот, не легко осуществить. Действительно, частота оптических колебаний огромна (порядка 10^{14}—10^{15} раз в секунду). Следовательно, для заметного изменения ее путем модуляции надо было бы иметь возможность изменять интенсивность световых волн сотни миллиардов раз в секунду, что, очевидно, крайне трудно выполнить. Правда, можно использовать чрезвычайную чувствительность оптических методов определения изменения частоты, которые позволяют открыть изменения периода, не превышающие одной десятитысячной процента. Благодаря этому можно было бы ограничиться модулированием световой волны в десятки миллионов раз в секунду, что, конечно, значительно проще. Хотя некоторые попытки такого рода и производились и, повидимому, не без успеха, тем не менее искусственное воспроизведение оптической модуляции чрезвычайно трудно и не может считаться надежно осуществленным. Слишком трудно при помощи экспериментальной техники получить достаточно быстрые изменения интенсивности. Вот почему кажется естественным обратиться к явлениям, имеющим место в природе, и постараться, таким образом, заполнить пробел в технике эксперимента.

Задача, подлежащая разрешению, напоминает до известной степени задачу создания условий, необходимых для реализации диффракции X-лучей. Принципиальная трудность в этом случае состояла в чрезвычайно малой длине волны X-лучей, так что для надежного диффракционного опыта требовалось создание щелей и экранов, размеры которых

были бы гораздо меньше обычно употребляющихся в оптике. Известно, сколько остроумия и изобретательности было затрачено при попытках создать искусственно условия, необходимые для наблюдения диффракции X-лучей, но все получавшиеся в этих опытах результаты оставались ненадежными и двусмысленными. Положение изменилось радикально, когда Лауэ выразил замечательную по своей простоте идею использовать естественную периодическую структуру кристалла вместо искусственных устройств. Этим моментом датируется надежное установление диффракции X-лучей и в то же время начинается плодотворная эпоха изучения кристаллического строения.

В проблеме, нас интересующей, надежное и плодотворное осуществление оптической модуляции точно так же оказалось возможным, благодаря использованию некоторых естественных процессов. Эти процессы связаны с явлениями рассеяния света и приводят нас, таким образом, вновь к нашей основной теме.

Мы уже упоминали, что первичный эффект воздействия света на вещество состоит в поляризации его молекул, т. е. в смещении электрических зарядов, их составляющих, благодаря чему молекула получает известный электрический момент. Показатель преломления вещества определяется значением электрического момента, возникающего под действием поля напряженностью в одну единицу и отнесенного к единице объема, т. е. способностью молекул к поляризации (поляризуемостью), и числом молекул в единице объема (плотностью вещества).

Для того чтобы среда была оптически однородна, показатель преломления ее во всех участках должен быть одним и тем же. Если показатель преломления меняется случайным образом от одного участка к другому, причем размеры этих участков сравнимы с длиной световой волны, то среда будет мутной, т. е. будет рассеивать свет. Итак, есть две возможные причины оптической неоднородности: нерегулярные изменения плотности и нерегулярные изменения поляризуемости. И та и другая могут обусловить рассеяние света. Обе причины дают себя знать в средах, которые называются мутными в обычном значении этого слова и которые образованы включением в практически однородное вещество (например, воду) мелких посторонних частичек (например, пылинок мела).

Каждая из этих причин может проявляться и в тех нарушениях однородности среды, которые обусловлены молекулярным строением

вещества и ведут к рассеянию света, называемому *молекулярным* рассеянием.

Первая из этих причин была рассмотрена, как я уже указывал, Смолуховским (1908), который ввел представление о флюктуации плотности. В среде, построенной из молекул, постоянство плотности в элементарных объемчиках означает, что молекулы характеризуются строго регулярным расположением. Но такое расположение невозможно. В самом деле, даже для идеального кристалла при абсолютном нуле температуры молекулярная система обладает некоторым запасом энергии, которая проявляется в непрерывных нарушениях строгой регулярности кристаллического строения. При повышении температуры эти запасы внутренней энергии увеличиваются и вместе с тем еще сильнее нарушается правильное расположение молекул. Под действием теплового движения молекулы в отдельных участках то сближаются, то удаляются, так что средняя плотность вещества испытывает кратковременные вариации — флюктуации плотности, которые и ведут к вариациям показателя преломления. Среда становится мутной.

Таков ход идей, принадлежавших в основном Смолуховскому. Можно продолжить развитие этих соображений в духе представлений о модуляции, о котором я упоминал выше.

Флюктуации плотности, представляющие собою случайное скопление или разрежение молекул, не устойчивы и должны рассасываться с течением времени. Таким образом, неоднородности среды, вызванные флюктуациями плотности, меняются со временем, а следовательно должна меняться во времени и интенсивность света, рассеянного на этих неоднородностях. Конечно, эти изменения интенсивности нельзя наблюдать в виде миганий, ибо они обусловливаются молекулярными процессами и вследствие этого происходят весьма быстро. Наши воспринимающие аппараты — глаз, фотоэлемент, фотопластинка — не могут регистрировать столь быстрых изменений интенсивности и воспринимают лишь некоторую среднюю интенсивность. Но на основании сказанного выше легко можно заключить, что эта вариация интенсивности должна проявляться в изменении частоты света, ибо она есть не что иное, как модуляция рассеянного света. Она происходит, вообще говоря, вполне беспорядочно. Однако теоретическое рассмотрение молекулярных процессов, обусловливающих указанное рассасывание флюктуаций, позволяет установить

важные закономерности общего характера. Наиболее ясные результаты получаются при рассмотрении явлений, происходящих в кристаллах. Такого рода рассмотрение было выполнено Л. И. Мандельштамом еще в 1918 г., хотя соответствующая заметка («О рассеянии света неоднородной средой») появилась значительно позже, в 1926 г., когда часть найденных Л. И. Мандельштамом результатов была уже опубликована Л. Бриллюэном (1922).

Главную роль при выравнивании уплотнений и разрежений играют упругие свойства кристалла. Правильная структура кристалла обусловливается значительными междумолекулярными силами, действующими между атомами, из которых построен кристалл. Всякое скопление или разрежение атомов, слегка искажающее правильную структуру кристалла, должно выправляться под действием междумолекулярных сил, подобно тому как всякое уплотнение в кристалле, вызванное толчком, не остается локализованным, а разбегается по кристаллу в виде упругой волны. Другими словами, всякое случайное скопление молекул, возникшее вследствие их теплового движения, дает начало упругой волне, разбегающейся по кристаллу.

Представление о таком механизме тепловых движений в кристалле было в 1912 г. успешно приложено Дебаем для расчета удельных теплот одноатомных кристаллов. Беспорядочная смена уплотнений и разрежений эквивалентна совокупности упругих волн различной частоты, пробегающих кристалл по всевозможным направлениям. Эти волны сообщают кристаллу пространственную неоднородность, пространственную «рябь», подобно тому как горсть песка, брошенная в спокойную воду, вызывает на поверхности рябь, покрывая ее сеткой волн различной частоты, разбегающихся по всем направлениям.

Свет, распространяющийся внутри деформированного таким образом кристалла, рассеивается во все стороны, причем в каждом из заданных направлений свет определенной длины волны рассеивается благодаря наличию одной из пространственных волн определенной частоты, составляющих всю совокупность упругих волн, рябящих внутренность кристалла. Упругие волны другого периода вызывают рассеяние света в другом направлении и т. д. Так осуществляется рассеяние по всем направлениям, которое и наблюдается экспериментально.

Таким образом, для рассеяния монохроматического света в одном

определенном направлении существенна упругая волна вполне определенной частоты. Под действием этой волны плотность кристалла испытывает периодические изменения определенной частоты, а следовательно и интенсивность рассеянного света подвергается периодической модуляции во времени с соответствующей частотой.

Анализ показывает, что частота этой модуляции значительна, достигая для кварца, например, порядка 10^{10} раз в секунду при собственной частоте света в $4 \cdot 10^{14}$ (фиолетовый свет). Эта модуляция должна вызывать изменение частоты рассеянного света, примерно, на одну трехсотую процента.

Такое изменение достаточно значительно, чтобы его можно было уверенно констатировать оптическими методами. Главное осложнение опыта заключается в чрезвычайной слабости рассеянного света, что крайне затрудняет применение аппаратов большой разрешающей силы, необходимых для установления малых изменений частоты. Тем не менее ожидаемый эффект модуляции нужного порядка величины был наблюдаем в опытах Л. И. Мандельштама и моих и еще лучше в опытах Е. Ф. Гросса, поставленных по нашему предложению.

Гроссу удалось более детально исследовать тонкую структуру линий рассеяния и подтвердить теорию Мандельштама и с количественной стороны. Более того, Гросс обнаружил существование резкой тонкой структуры при рассеянии света также и в жидкостях — важный результат, который показал, что описанный механизм упругого рассасывания флюктуации применим и к жидкостям, где теоретически можно было опасаться слишком большого затухания упругих волн, которое могло бы смазать описываемое явление.

Вторая причина, способная вызвать нарушение оптической однородности среды и обусловить рассеяние света, состоит в нерегулярном изменении поляризуемости молекул в элементарных объемчиках, составляющих среду. Молекулы построены, как известно, из электрических частиц двух родов: легких электронов и массивных ионов. Так как обычный свет, видимый или ультрафиолетовый, представляет собою электромагнитное поле высокой частоты (10^{14}—10^{15}), то он может воздействовать непосредственно только на легкие электроны, тогда как для воздействия на ионы требовалось бы поле несравненно более низкой частоты, порядка 10^{12} раз в секунду, т. е. инфракрасный свет.

Тем не менее мы отнюдь не можем не учитывать влияния ионов. Наоборот, взаимное расположение и характер ионов, в поле которых движутся электроны, как раз и определяют, насколько легко могут смещаться электроны под действием постороннего поля, например, поля световой волны, т. е. определяют поляризуемость молекулы, а следовательно и показатель преломления среды.

Представим себе, что ионы колеблются друг относительно друга с частотой, свойственной молекуле (т. е. около 10^{12}), и связанное с этим изменение в конфигурации зарядов вызывает изменение поляризуемости, а следовательно и изменение показателя преломления в данном участке среды. Так как в разных участках среды могут, конечно, иметь место некоторые вариации указанного изменения показателя преломления, то оптическая однородность нарушится: среда становится мутной и будет рассеивать свет.

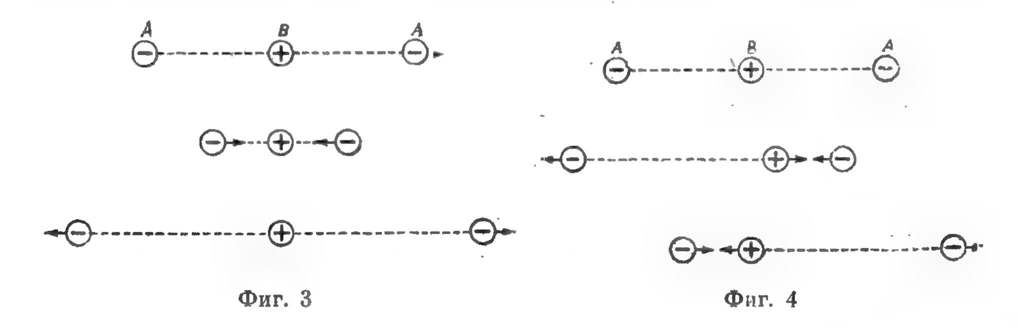
Этот новый вид рассеяния связан, как мы видим, с изменениями показателя преломления, обусловленными колебаниями ионов и, следовательно, происходящими в темпе этих колебаний, т. е. в темпе собственных частот молекулы (имеющих порядок 10^{12} колебаний в секунду). Поэтому при этом темпе рассеяния свет будет модулироваться с частотой, гораздо более высокой, чем в процессах, рассмотренных выше и связанных с рассасыванием флюктуации плотности.

Вызываемое этим новым видом модуляции изменение частоты света будет весьма значительным, так что если падающий свет имеет частоту ν, то в свете рассеянном появляются также частоты $(\nu \pm m)$, где m — частоты собственных колебаний молекул. Этот вид рассеяния характеризуется, следовательно, частотами, представляющими собою комбинацию из частоты падающего света ν и частоты собственного колебания молекулы m, откуда и произошло название — комбинационное рассеяние света, предложенное нами для этого явления.

Изложенные соображения разъясняют также, что не всякое собственное колебание молекулы может стать причиной рассеяния света и проявиться в спектре комбинационного рассеяния. Для этого необходимо, чтобы колебание ионов меняло периодически поляризуемость молекулы, т. е. сопровождалось таким изменением конфигурации электрических зарядов молекулы, при которой поле внутри молекулы за время первой половины колебания отличалось бы от поля, соответствующего второй

половине. Однако это не всегда может иметь место. Простой пример поясняет положение дела.

Предположим, что мы имеем молекулу, состоящую из трех заряженных атомов, расположенных на прямой линии, например молекулу $A - B - A$ (фиг. 3 и 4). Эта молекула может совершать два типа колебаний. При первом (фиг. 3) отрицательные заряды то приближаются, то удаляются от положительного, который остается на месте; при этом



Фиг. 3 Фиг. 4

периодически меняется поле внутри молекулы и, следовательно, ее поляризуемость. Такое колебание модулирует рассеянный свет и проявляется в спектре комбинационного рассеяния света в виде интенсивной линии.

Но это колебание не может быть вызвано приходящей инфракрасной волной, ибо электрическое поле этой волны не может двигать два отрицательных заряда в разные стороны и оставлять положительный на месте. Поэтому, если на такую молекулу падает инфракрасный свет подходящей частоты, он не будет ею поглощаться: в инфракрасном спектре поглощения такая частота не обнаружится.

Наоборот, колебание типа, изображенного на фиг. 4, должно соответствовать сильной полосе поглощения, ибо под действием поля световой волны подходящей частоты заряды, составляющие нашу молекулу, придут в сильное колебание типа, изображенного на фиг. 4, на что будет затрачена энергия приходящей волны: вещество будет хорошо поглощать инфракрасный свет соответствующей частоты. Но это колебание не будет модулировать рассеянного света, ибо при этом типе колебания поляризуемость молекулы практически не меняется: две крайние конфигурации зарядов (фиг. 4) совершенно идентичны, что легко видеть, если перевер-

нуть рисунок так, чтобы нижняя часть его оказалась наверху. Но если при колебании поляризуемость молекулы не меняется, то такое колебание не проявится в спектре комбинационного рассеяния.

Таким образом, становится понятным, что могут быть колебания, дающие себя знать как линии комбинационного рассеяния и не проявляющиеся в инфракрасных спектрах поглощения, и наоборот. Конечно, могут быть колебания, активные как в спектре рассеяния, так и в спектре поглощения. Эти рассуждения устраняют все сомнения, которые может пробудить рассмотрение табл. 2, и показывают в то же время, что инфракрасные спектры могут иногда с пользой дополнять данные, полученные по методу комбинационного рассеяния.

Итак, мы видим, что наряду с рассеянием, обусловленным флюктуациями плотности и характеризующимся тонкой структурой рассеянного света, трудно поддающейся наблюдению, должно существовать и рассеяние, обусловленное флюктуациями поляризуемости молекул, проявляющееся как эффект модуляции рассеянного света собственными колебаниями молекул.

Для кристаллов этот новый вид рассеяния представляет собой вполне естественное дополнение к рассеянию, обусловленному флюктуациями плотности, которые, как мы видели, могут быть представлены как совокупность дебаевских упругих волн. Дебаевская теория удельных теплот имела в виду кристаллы, построенные из атомов (одноатомные), к которым относится, например, большинство металлов. В более полной теории удельных теплот, развитой Борном и Карманом для многоатомных кристаллов, учитываются кроме дебаевских (акустических) колебаний также и взаимные колебания различных атомов, входящих в элементарную ячейку многоатомного кристалла. Колебания эти происходят с частотами, характерными для данного кристалла, и носят название оптической или инфракрасной части упругих колебаний. Таким образом, существование комбинационного рассеяния света, представляющее собой модуляцию рассеянного света инфракрасными колебаниями молекул, закономерно вытекает из идей Л. И. Мандельштама о роли модуляции в оптике, сформулированной им, как уже указывалось, еще в 1918 г.

Закономерно также и то, что явление комбинационного рассеяния света было обнаружено Мандельштамом и мной в ходе экспериментальных исследований по тонкой структуре рассеянного света. Однако явле-

ние не было нами предусмотрено заранее, и лишь после его установления нам сделалась ясной связь, изложенная выше.

Исторически дело сложилось так. Когда Л. И. Мандельштам занял в 1925 г. кафедру в Московском университете, он выдвинул в качестве первой экспериментальной проблемы по оптике задачу об исследовании тонкой структуры рассеянного света в кристаллах. В 1925 г. самый факт молекулярного рассеяния света в кристаллических телах не был еще установлен достаточно надежно. Поэтому необходимо было начать изучение рассеяния света в кристаллах с самого начала. Только когда мною была решена первая задача — установление молекулярного характера наблюдаемого рассеяния и найдены критерии, позволяющие отличать молекулярное рассеяние от рассеяния на случайных включениях в кристалле, — оказалось возможным приступить в 1927 г. к основной задаче.

Отсутствие интерференционных спектральных аппаратов достаточной разрешающей силы заставляло искать обходных решений. Была предпринята попытка установить изменение длины волны рассеиваемого света методом резонансного поглощения, состоящего в том, что в качестве первичного света применяется свет резонансной линии ртути (λ 2537 Å), а рассеянный свет направляется в обычный призматический спектрограф сквозь фильтр из паров ртути. Таким образом, спектрограф небольшой разрешающей силы, примененный в этих опытах, не предназначался, конечно, к наблюдению ничтожного изменения длины волны, составляющего тонкую структуру рассеянного света, а должен был играть вспомогательную роль.

Однако полученные в процессе исследования спектрограммы обнаружили слабые следы спектральной линии, характеризующейся значительно бо́льшим изменением длины волны, чем ожидаемое по теории. Тщательные и весьма разнообразные проверки показали реальность наблюдаемого явления, установив с полной несомненностью наличие в рассеянном свете длин волн, заметно отличающихся от первичной длины волны. Стало ясным, что наблюдаемое в этих опытах изменение длины волны есть результат модуляции рассеянного света другим процессом, более быстрым, чем процесс упругого рассасывания уплотнений. Были экспериментально установлены основные законы явления, выяснена связь наблюдаемого изменения частоты с собственными инфракрасными

частотами вещества и, таким образом, открыта причина нового вида модуляции. В то же время дана была и квантовая интерпретация нового явления, необходимая для понимания вопроса о соотношении интенсивности красных и фиолетовых сателлитов, интерпретация, подтвержденная специальными опытами по влиянию температуры, описанными в начале этой статьи.

Наше изложение показывает, что классическое (релеевское) рассеяние света есть проявление междумолекулярных сил, ведущих к образованию и рассасыванию флюктуаций плотности и вызывающих модуляцию рассеянного света, проявляющуюся в тонкой структуре линий рассеянного света; явление же комбинационного рассеяния света есть проявление модуляции рассеянного света, вызываемой процессами, протекающими в самой молекуле, а именно колебаниями, обусловленными внутримолекулярными силами. Поэтому свет комбинационного рассеяния несет на себе печать внутримолекулярных процессов.

Изучая спектральные особенности рассеянного света, мы получаем возможность реконструировать модулирующие факторы, т. е. изучать внутримолекулярный механизм, вызвавший эту модуляцию. Подобно тому как, принимая модулированные радиотелефонные сигналы, мы слышим голос говорящего перед микрофоном, так, наблюдая спектр комбинационного рассеяния света, мы, так сказать, слушаем рассказ молекулы об особенностях происходящих в ней процессов.

Если молекула, подобная молекуле CO_2, может характеризоваться двумя типами колебаний, изображенных на фиг. 3 и 4, то первое колебание не вызывает модуляции и не проявляет в спектре комбинационного рассеяния света, а второе отчетливо заявляет о себе появлением на спектрограмме соответствующего сателлита. Мы, следовательно, не только можем установить частоту этого колебания, но и сделать далеко идущие заключения об особенностях этого колебания.

При решении таких более специальных вопросов интерпретация явления облегчается изучением интенсивности сателлитов и других его особенностей (в частности, характера поляризации света сателлита), сопоставлением с данными инфракрасной абсорбции и теоретическими соображениями совершенно общего характера, касающимися симметрии молекулы. В результате структурные особенности модели выступают нередко с полной ясностью, открывается возможность проверить и отвергнуть

те или иные гипотезы, относящиеся к структуре молекулы, и разрешить сложные или сомнительные пункты. Конечно, принимая во внимание сложность и многообразие относящихся сюда проблем, мы должны признать, что мы еще далеки от полного их разрешения; однако в целом ряде случаев достигнуты поразительные успехи. Плодотворность описываемого метода и относительная экспериментальная легкость его применения объясняют энергичную работу, которая проводится в этом отношении во всех лабораториях мира. За двадцать лет, истекающих со времени открытия комбинационного рассеяния света, опубликовано уже около трех тысяч работ, большинство которых представляют экспериментальные исследования.

В небольшой статье нет никакой возможности дать представление об огромном комплексе вопросов из области теоретической химии, главным образом органической, которые получили успешное разрешение благодаря применению метода комбинационного рассеяния. Я ограничусь поэтому примерами лишь из двух областей, которые разрабатывались главным образом в нашей лаборатории за последние годы.

Как уже указывалось, в спектре комбинационного рассеяния запечатлеваются особенности молекулярной динамики, т. е. в первую очередь проявляются силы внутримолекулярные, обусловливающие частоту и форму собственного колебания молекулы. Однако при исследовании конденсированных систем или газов при высоком давлении мы не можем отвлечься и от действия сил междумолекулярных, которые могут наложить свой отпечаток на поведение молекулы.

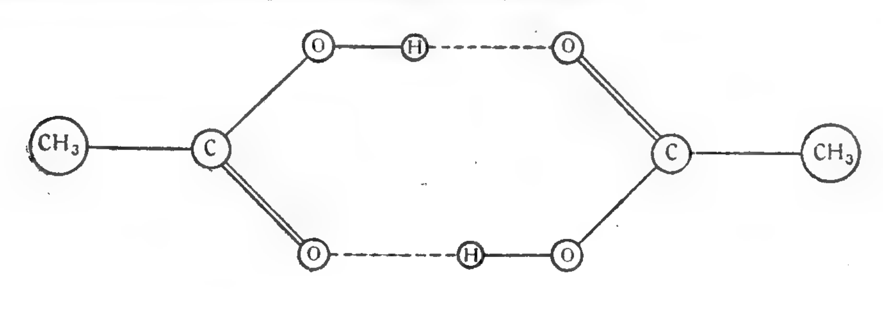
Так как, вообще говоря, междумолекулярные силы значительно слабее внутримолекулярных, то действие их должно сказаться лишь на второстепенных деталях явления. И действительно, в большинстве случаев спектр рассеяния вещества в газообразной фазе практически не отличается от спектра того же вещества в жидком или кристаллическом состоянии.

Есть, однако, один вид междумолекулярного взаимодействия, который занимает промежуточное положение между внутримолекулярными силами (силами химической связи) и обычными ван-дер-ваальсовыми силами, проявляющимися между молекулами и обусловливающими переход из одного агрегатного состояния в другое. Этот вид междумолекулярных сил получил название сил *водородной* связи; они проявляются

23*

отчетливее всего в том случае, когда в состав одной из молекул входит группа O — H (гидроксил), а в состав соседней — атом кислорода (а также фтора или азота).

Водородная связь ведет к образованию агрегатов типа R—O—H...O——R′, причем положение водорода между двумя кислородами обусловливает сильное взаимодействие между соседними молекулами, — взаимодействие по преимуществу электростатического характера.

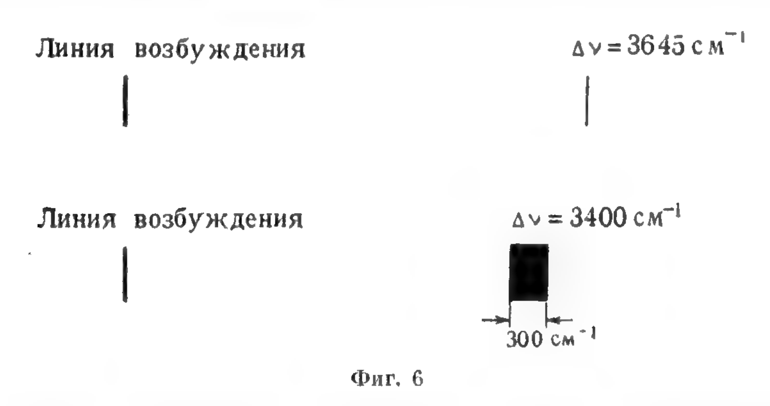


Фиг. 5

Во многих физико-химических проблемах эта водородная связь проявляется как источник сильного междумолекулярного взаимодействия. Ею объясняется способность к ассоциации таких жидкостей, как вода или спирты, характеризующиеся значительной теплотой испарения, ибо при испарении необходимо разорвать сильные водородные связи; водородная связь привлекается для объяснения образования парных молекул некоторых жирных кислот по типу фиг. 5, где две водородные связи соединяют две молекулы в прочный димер, и т. д.

Естественно ожидать, что такое сильное междумолекулярное взаимодействие должно проявляться и в спектрах комбинационного рассеяния и, следовательно, эти последние могут быть привлечены для изучения вида междумолекулярных взаимодействий. Действительно, уже в первых опытах по изучению спектра рассеяния водяного пара и воды обнаружилось проявление водородной связи. В то время как спектр паров воды характеризуется резкой линией комбинационного рассеяния, соответствующей частоте колебания группы O — H ($\Delta\nu = 3645$ см$^{-1}$), спектр жидкой воды дает вместо этой линии размытую полосу, положение центра которой соответствует частоте колебаний примерно $\Delta\nu = 3400$ см$^{-1}$.

На фиг. 6 дано схематическое представление об указанном различии спектров. Это явление было изучено в многочисленных работах нашей лаборатории (М. И. Батуев, Ф. С. Барышанская, В. И. Малышев, С. А. Ухолин, И. А. Яковлев и др.). Были исследованы вода и спирты в широком интервале температур и давлений, от парообразного состояния до твердого, включая и критическое состояние. Были изучены многочисленные вещества, содержащие O — H-группу (вода, длинный гомологический ряд спиртов, фенолы, жирные кислоты, гидроокиси металлов, кристаллогидраты, содержащие кристаллизационную воду, и т. д.),

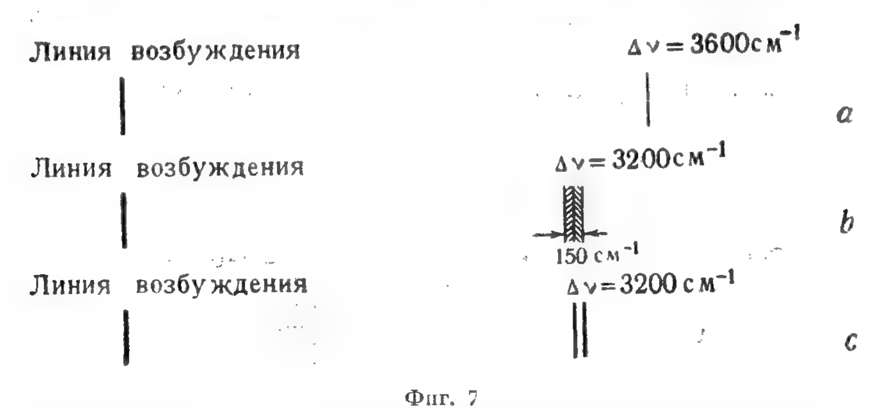


Фиг. 6

причем широко варьировалась температура, изучались растворы в разнообразных растворителях в широком интервале концентраций и исследовались кристаллические вещества разнообразной кристаллографической структуры.

Как общий итог всех исследований можно считать установленным наличие спектроскопического признака водородной связи, состоящего в уменьшении собственной частоты O — H-колебания под влиянием возмущающего действия соседнего кислорода (или азота). Увеличение взаимного расстояния между соседними молекулами (путем уменьшения плотности при нагревании, растворения в нейтральных растворителях, перехода к соответствующим кристаллическим формам) ведет к разрыву водородной связи и восстановлению собственного колебания O — H-группы со свойственной ей частотой (около 3600 $см^{-1}$). Опыты с кристаллами позволили даже установить, что таким критическим расстоянием, при

котором возмущающее действие соседнего кислорода не сказывается (водородная связь не образуется), является расстояние около 2.8Å, т. е. двойной радиус кислорода. Расширение линии в полосу, сопровождающее обычно образование водородной связи, представляет собою вторичное явление, обусловленное вариациями взаимного расстояния между О — Н-группой и возмущающим ее кислородом. Эти вариации происходят вследствие теплового движения атомов и уменьшаются при значительном понижении температуры. Так, например, исследование спектра



Фиг. 7

a — невозмущенное О — Н-колебание (в парах); *b* — возмущенное О — Н-колебание (кристалл. $B(OH)_3$) при $T = 400°$ K; *c* — возмущенное ОН-колебание (кристалл. $B(OH)_3$) при $T = 100°$ K.

$B(OH)_3$ показывает, что при температуре около 400° K мы имеем сильно смещенную по сравнению с колебанием невозмущенной О — Н-группы и расширенную полосу (с частотой около 3200 $см^{-1}$ при ширине около 150 $см^{-1}$); при температуре же 100° K полоса превращается в узкую линию, хотя смещение ее полностью сохраняется. Фиг. 7 дает представление о характере явлений; $B(OH)_3$ характеризуется двумя близкими линиями (фиг. 7, *c*), которые при нагреве дают сливающиеся полосы (фиг. 7, *b*).

Таким образом, явление комбинационного рассеяния оказалось чрезвычайно плодотворным при исследовании физико-химической проблемы водородной связи. Другая услуга, которую оказывают явления комбинационного рассеяния света, относится к химии аналитической. Спектр комбинационного рассеяния представляет настоящий спектральный пас-

порт молекулы или, скорее, ее дактилоскопический снимок. Эти спектры могут служить для идентификации молекул, подобно тому как эмиссионные спектры атомов служат со времен Кирхгофа и Бунзена для идентификации атомов. Поэтому на методе комбинационного рассеяния может быть основан настоящий метод качественного и количественного спектрального анализа молекул. И в этом отношении метод особенно полезен для органической химии, ибо он позволяет различать многочисленные изомерные соединения, для анализа которых химия зачастую не имеет нужных методов.

В руководимой мною лаборатории проделана обширная работа по разработке методов качественного и количественного анализа смесей органических соединений по методу комбинационного рассеяния света (М. И. Батуев, П. А. Бажулин, А. С. Елинер, Ю. В. Розенберг, Х. Е. Стерин, М. М. Сущинский и др.). Накоплен значительный материал, характеризующий спектры почти сотни индивидуальных углеводородов, в том числе всех изомеров парафинового ряда, кончая октанами, причем исследованы не только частоты всех линий этих углеводородов, но и тщательно измерены интенсивности большинства этих линий, а для некоторых определена также степень деполяризации.

Эти данные позволяют уже теперь проводить качественные и количественные анализы смесей, содержащие углеводороды, с температурой кипения до 130° С. Они были применены к анализу авиационных бензинов (отечественных, импортных и трофейных) и к решению ряда конкретных аналитических задач, связанных с переработкой моторного топлива (анализ продуктов крекинга, гидрогенизации и т. д.). По мере обогащения составляемого нами атласа спектров индивидуальных углеводородов [1] аналитические приложения метода будут расширяться и совершенствоваться.

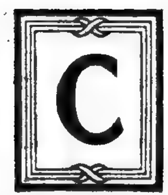
На этих двух примерах из практики нашей лаборатории видно, какие широкие перспективы имеет метод комбинационного рассеяния в области всевозможных приложений.

---

[1] Богатая коллекция индивидуальных углеводородов, которой мы в настоящее время располагаем, составлена из веществ, тщательно синтезированных в лаборатории академика Б. А. Казанского [Институт органической химии АН СССР].

Академик
А. Н. Теренин
и член-корреспондент АН СССР
С. Э. Фриш

РАБОТЫ Д. С. РОЖДЕСТВЕНСКОГО И ЕГО ШКОЛЫ ПО АНОМАЛЬНОЙ ДИСПЕРСИИ В ПАРАХ МЕТАЛЛОВ



Самую обширную и значительную часть многообразного научного наследия покойного академика Дмитрия Сергеевича Рождественского составляют его исследования по аномальной дисперсии в парах металлов. В течение всей своей научной деятельности он много раз возвращался к теме об аномальной дисперсии: с нее он начал, готовясь еще молодым человеком к магистерской диссертации, ей была посвящена его последняя работа, опубликованная уже после его смерти (1940).

Особенный размах работы по аномальной дисперсии получили в годы советской власти, когда под руководством Д. С. Рождественского в Государственном Оптическом институте и в Ленинградском университете стали работать его ученики и сотрудники. Рождественским был создан совершенно новый и оригинальный метод точного количественного исследования аномальной дисперсии, получивший название метода «крюков». При помощи этого метода, ставшего ныне классическим, Д. С. Рождественский и его ученики поставили и решили большое число задач, имеющих первостепенную важность для теории строения атомов. Изучение аномальной дисперсии стало монополией советских физиков. Немногочисленные работы по аномальной дисперсии, выполненные за границей,



Д. Рождественский.

произведены методом «крюков» с помощью аппаратуры, скопированной с аппаратуры Рождественского.

Прошло 30 с лишком лет с момента первой работы Рождественского, однако открытый им метод не потерял своей свежести — случай очень редкий при современных бурных темпах развития физики. И в настоящее время целый ряд актуальных задач экспериментальной и теоретической спектроскопии решается методом «крюков» Рождественского.

Работы Рождественского и его учеников, если их рассматривать в исторической последовательности, ясно отражают все этапы в развитии учения о строении атомов.

В период развития классической электронной теории казалось несомненным, что электроны внутри атомов удерживаются квази-упругими силами. Такая гипотеза позволила решить ряд важных задач, в частности задачу о резонансном раскачивании электронов падающей световой волной, что дало возможность объяснить явления аномальной дисперсии. Свет, проходя через прозрачное вещество, преломляется в нем, причем свет разного цвета, т. е. разных частот, преломляется по-разному. Другими словами, показатель преломления n есть некоторая функция от частоты света ν. По формуле, впервые данной Зельмейером и получившей обоснование в теории квази-упругих электронов, показатель преломления n равен

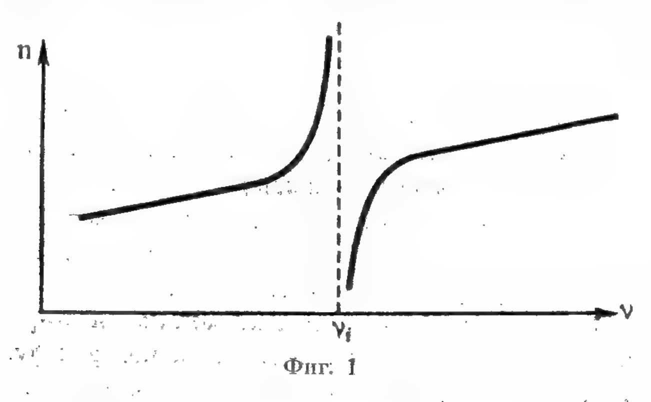
$$n=1+\sum_i \frac{a_i}{\nu_i^2-\nu^2}, \qquad (1)$$

где ν_i — частота собственных колебаний электронов внутри атома. Величина a_i представляет собою константу, связанную с зарядом и массой электрона e и m, а также с числом электронов N_i, обладающих собственной частотой колебаний ν_i; эта связь выражается соотношением

$$a_i=\frac{e^2}{2\pi m}N_i. \qquad (2)$$

Формула Зельмейера (1) выражает резонансный характер взаимодействия падающего света с атомами. В спектре вблизи линии поглощения показатель преломления сильно возрастает по одну сторону от линии и круто спадает по другую сторону от нее. Этот ход показателя преломления n схематически изображен на фиг. 1. Он получил название аномальной дисперсии, хотя ничего «аномального» в нем нет; наоборот, он соот-

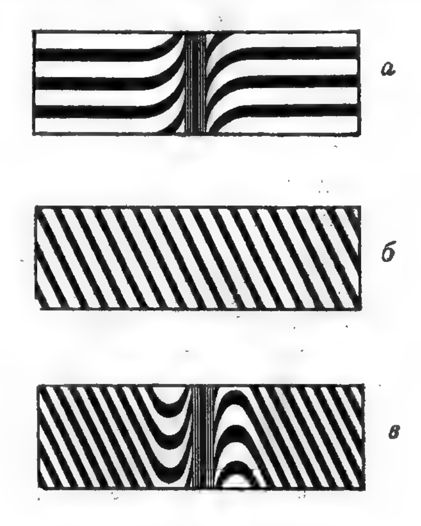
ветствует физической сущности взаимодействия между атомами и светом. Экспериментальное изучение аномальной дисперсии было интересно в двух отношениях: 1) для проверки гипотезы о квази-упругом характере внутриатомных сил и 2) для определения числа электронов N_i, обладающих собственной частотой колебаний ν_i. Эти электроны определяют одновременно и ход аномальной дисперсии и интенсивность линий поглощения; их число — существенная характеристика строения атомов.



Фиг. 1

В то время, когда Д. С. Рождественский приступил к своей первой работе, сколько-нибудь точного метода изучения аномальной дисперсии не было. Вуд использовал для наблюдения аномальной дисперсии старый способ «скрещенных призм», впервые предложенный еще Ньютоном для изучения дисперсии в стеклах. Однако способ Вуда не давал возможности производить надежные измерения. Более обещающим был метод скрещенных интерферометра и спектрографа, предложенный Пуччианти. Этот метод вначале применил и Рождественский, придя, кстати сказать, к нему самостоятельно и только впоследствии ознакомившись с работой Пуччианти. Сущность метода Пуччианти — Рождественского заключается в следующем. Интерференционные полосы, получающиеся в белом свете, проектируются на вертикальную щель спектрографа. Положение максимума, соответствующего нулевой разности хода, не зависит от частоты света. Отсюда этому максимуму в сплошном спектре, даваемом спектрографом, отвечает горизонтальная светлая полоса, тянущаяся вдоль всего спектра. Максимумы же, соответствующие разностям хода в одну, две, три и т. д. длины волны, располагаются для разных частот падающего

света на разных высотах щели. Им отвечают в сплошном спектре светлые полосы, расположенные выше и ниже нулевой полосы и расходящиеся веером от синей части спектра к красной. Между светлыми полосами располагаются темные, отвечающие минимумам в интерференционной картине, отброшенной на щель. В результате сплошной спектр оказывается вдоль всей длины прочерченным попеременно светлыми и темными полосами. Если ввести в интерферометр на пути одного из лучей исследуемое вещество в виде плоскопараллельного однородного столба паров, то полосы сместятся на величину, пропорциональную введенной разности хода. Разность же хода при данной толщине столба паров пропорциональна $n - 1$, где n — показатель преломления паров.



Фиг. 2

Таким образом, смещение интерференционных полос в спектрографе непосредственно воспроизводит ход дисперсии во всей наблюдаемой области. Если в эту область попадет какая-нибудь из линий поглощения, то интерференционные полосы загнутся около нее, как показано на фиг. 2, *а*. Их изгиб весьма эффектно и наглядно изобразит ход аномальной дисперсии вблизи линии поглощения.

Однако и этот метод изучения аномальной дисперсии мало пригоден для получения количественных данных; он требует промера большого числа точек на круто восходящей или ниспадающей части интерференционной полосы, где измерения кропотливы и недостаточно точны. Вблизи самой линии поглощения полосы идут так круто и расположены так близко друг к другу, что промеры вообще становятся невозможными.

Д. С. Рождественский подверг способ Пуччианти существенному и в высшей степени остроумному изменению. Во вторую ветвь интерферометра он ввел плоскопараллельную стеклянную пластинку определенной толщины. При отсутствии паров в первой ветви интерферометра введение такой пластинки дало бы смещение полос, изображающее в ней ход дис-

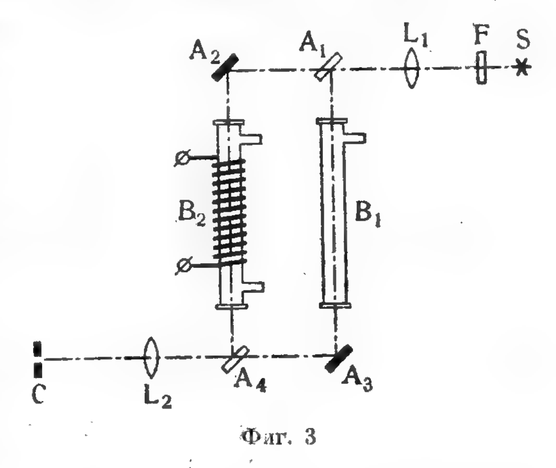
персии. В прозрачной стеклянной пластинке, вдали от полос поглощения, ход дисперсии сводится к монотонному убыванию коэффициента преломления с возрастанием длины волны («нормальная» дисперсия), что ведет лишь к наклону полос. При большой толщине пластинки наклон значителен. Наконец, так как пластинка введена в другую ветвь интерферометра, не в ту, где пары, то наклон полос получается в обратную сторону. Вид этих полос изображен на фиг. 2, *б*.

При одновременном действии паров с аномальной дисперсией и стеклянной пластинки с нормальным ходом дисперсии смещение полос определится алгебраической суммой смещений, вызванных парами и стеклянной пластинкой в отдельности. Вдали от полосы поглощения паров, где их показатель преломления мало зависит от длины волны, сильнее выражен наклон, вызванный стеклянной пластинкой. Вблизи полосы поглощения, где изгиб интерференционных полос, вызванный дисперсией паров, очень велик, сильнее изгиб, даваемый парами. В результате полосы в некотором месте перегнутся и дадут своеобразные максимумы и минимумы, схематически представленные на фиг. 2, *в*.

Эти изгибы Рождественский и назвал «крюками». При данной толщине стеклянной пластинки положение крюков определится величиной дисперсии около рассматриваемой линии поглощения паров. Чем сильнее дисперсия паров вблизи линии, тем дальше от линии образуются крюки. Таким образом, по положению крюков можно непосредственно измерить величину дисперсии паров около линии поглощения. Подбором соответствующей толщины стекла крюки всегда можно отодвинуть от линии и добиться такого их положения, при котором измерения будут легко выполнимыми и точными. В этом огромное преимущество метода Рождественского: измерения производятся просто, точно и могут выполняться на столь малых расстояниях от линий поглощения, где все другие методы перестают быть пригодными.

На фиг. 3 представлена в несколько упрощенном виде схема установки Рождественского: S — вольтова дуга, являющаяся мощным источником сплошного спектра, F — светофильтр с водой для поглощения инфракрасных лучей, L_1 — линза, дающая параллельный пучок лучей. Лучи падают на интерферометр, состоящий из четырех зеркал A_1, A_2, A_3, A_4, два из которых, A_2 и A_3, сплошные, остальные два полупрозрачные. Интерферометр представляет собою видоизменение интерферометра Жа-

мена и позволяет сильно развести друг от друга пучки света A_1A_3 и A_2A_4. В оба пучка вводятся одинаковые трубки B_1 и B_2, к концам которых прикреплены одинаковые плоскопараллельные стеклянные пластинки. Таким образом, трубки B_1 и B_2 сами по себе не вносят в пучки света добавочной разности хода. Пройдя через трубки и отразившись от зеркал A_3 и A_4, пучки света соединяются вместе и отображаются линзой L_2 на щель спектрографа C. В качестве спектрографа Д. С. Рождественский употреблял плоскую диффракционную решетку в автоколлимационной установке. Дисперсия решетки во втором порядке, в котором производились снимки, достигала 5 Å/мм.



Фиг. 3

Для наблюдения аномальной дисперсии в одну из трубок, например трубку B_2, вводился исследуемый металл. Трубка нагревалась электрической печью до температуры, достаточной для того, чтобы в ней получились пары при необходимом давлении. С целью образования «крюков» в пучок A_1A_3 вводилась плоскопараллельная стеклянная пластинка.

Для первых своих исследований Рождественский выбрал пары натрия, как наиболее удобные с экспериментальной точки зрения: они не требуют высокой температуры и имеют в видимой (желтой) части спектра две очень интенсивные линии поглощения. Эти линии обозначаются буквами D_1 и D_2; они располагаются близко друг к другу, но при помощи метода Рождественского можно промерить дисперсию не только вблизи каждой из этих линий, но и между ними.

Прежде всего Рождественский промерил точно ход кривой аномальной дисперсии вблизи D-линий натрия. Было промерено до 300 точек. Результаты измерений он сравнил с данными формулы Зельмейера (1); получилось прекрасное совпадение. Незначительные отступления наблюдались лишь на очень малых расстояниях от линии, при этом можно было показать, что эти отступления носят в известном смысле случайный,

«инструментальный» характер: при фотографировании спектра в разных порядках диффракционной решетки они получались разными. Истинное происхождение этих отступлений было выяснено лишь совсем недавно (1941) Г. С. Кватером, учеником и продолжателем работ Рождественского. Кватер выяснил, что причина отступлений лежит в характере распределения интенсивности в пределах интерференционной полосы; он указал критерии, при соблюдении которых отступления незаметны.

Таким образом, уже первая работа Рождественского показала весьма точную выполнимость формулы Зельмейера и, как казалось, подтвердила гипотезу о квази-упругом характере внутриатомных сил. Значительно позже, на сессии Академии Наук СССР в марте 1936 г., Рождественский сказал: «Вряд ли кто-нибудь тщательнее меня вдоль и поперек исследовал эти кривые, и мой ответ: да, эти кривые — гиперболы и, следовательно, электрон имеет квази-упругую связь. К счастью, присущий мне скепсис заставил меня прибавить: впрочем, достаточно линейности уравнений, заведующих колебаниями электрона, чтобы обусловить гиперболу, не входя в рассмотрение вопроса о существе связи». Последняя фраза весьма замечательна: она характеризует научную осторожность и прозорливость Рождественского. Он допускал возможность и иного толкования законов аномальной дисперсии, отличного от толкования, основанного на наглядном модельном представлении о квази-упругих колебаниях электрона внутри атома. Но если теоретические выводы и не были однозначны, то во всяком случае после работы Д. С. Рождественского было ясно одно: всякая будущая теория должна считаться с точной выполнимостью формулы Зельмейера.

В этой же работе Рождественский получил второй не менее важный результат. Он промерил отношение числа вибраторов N_1 и N_2 для D-линий натрия и получил для него значение 2 с точностью до 2.5%. Он показал, что это отношение не зависит ни от температуры, ни от давления паров и, следовательно, определяется свойствами самих атомов натрия. Несколько позже отношение, близкое к 2, он получил и для составляющих головных дублетов калия, рубидия и цезия. Таким образом, для отношения интенсивностей двух спектральных линий получилось целочисленное отношение. Этот факт, в то время совершенно непонятный, указал на своеобразие законов, которым подчиняются внутриатомные процессы; он нашел впоследствии объяснение лишь в квантовой механике.

Метод Рождественского дал простой и точный способ сравнения интенсивностей линий поглощения. Интенсивность линий поглощения определяется числом поглощающих центров N_i, которое по формуле (2) связано с константой a_i формулы Зельмейера. Вблизи линии поглощения в формуле Зельмейера (1) существенную роль играет лишь один член суммы. Поэтому для показателя преломления n вблизи линии поглощения приближенно можно написать

$$n-1=\frac{a_i}{\nu_i^2-\nu^2}.$$

Теория метода «крюков» показывает, что значение константы a_i получается простым измерением длин волн, при которых расположены «крюки». Такая простота измерений делает метод «крюков» наиболее удобным, быстрым и точным для определения интенсивностей спектральных линий поглощения. Другие методы — метод фотографической фотометрии, метод резонансного поглощения и метод вращения плоскости поляризации — более сложны и менее надежны.

Целочисленный характер отношения интенсивностей спектральных линий, обнаруженный Рождественским, впервые нашел отражение в теории атома Бора. По классической теории совокупность вибраторов с собственной частотой колебаний ν_i поглощает энергию

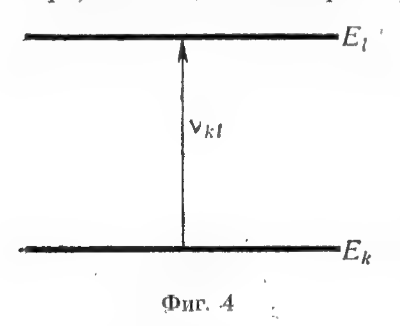
$$\Delta E=N_i\frac{\pi e^2}{m}\rho(\nu_i), \tag{3}$$

где N_i, e и m имеют те же значения, что и в формуле (2), а $\rho(\nu_i)$ — плотность энергии в падающем световом потоке частоты ν_i.

Теория Бора рассматривает поглощение и испускание света атомом как результат перехода атомной системы из одного дискретного энергетического состояния в другое. Если E_k — уровень более бедный энергией, а E_l — уровень более богатый энергией, то акту поглощения света соответствует переход с уровня E_k на уровень E_l, изображенный стрелкой на фиг. 4. Частота поглощения света ν_{kl} определяется правилом частот Бора

$$\nu_{kl}=\frac{E_l}{h}-\frac{E_k}{h},$$

где h — постоянная Планка. Если поглощающий пар содержит в единице объема N_k атомов, находящихся в состоянии с энергией E_k, то, по теории Бора, поглощенная энергия равна



Фиг. 4

$$\Delta E = N_k \frac{g_l}{g_k} \frac{c^3}{8\pi\nu_{kl}^2} A_{kl} \rho(\nu_{kl}), \tag{4}$$

где c — скорость света, а A_{kl}, g_k и g_l — константы, определяемые свойствами атома.

Вводя в рассмотрение величину

$$f_{kl} = \frac{g_l}{g_k} \frac{mc^3}{8\pi^2 e^2 \nu_{kl}} A_{kl}, \tag{5}$$

перепишем формулу (4) в виде

$$\Delta E = N_k f_{kl} \frac{\pi e^2}{m} \rho(\nu_{kl}). \tag{4a}$$

Она отличается от классической формулы (3) только тем, что, вместо числа гипотетических вибраторов N_l, стоит произведение из числа атомов N_k и некоторой константы f_{kl}, зависящей от свойств атомов. Это означает, что с точки зрения теории Бора поглощать могут все атомы, но переходы между разными энергетическими уровнями, а следовательно и соответствующие им акты поглощения, происходят не одинаково часто. Величина f_{kl} определяет вероятность данного перехода. Эти «вероятности перехода» f_{kl} непосредственно определяются методом «крюков» Рождественского. В самом деле, равенство

$$N_l = N_k f_{kl},$$

дает

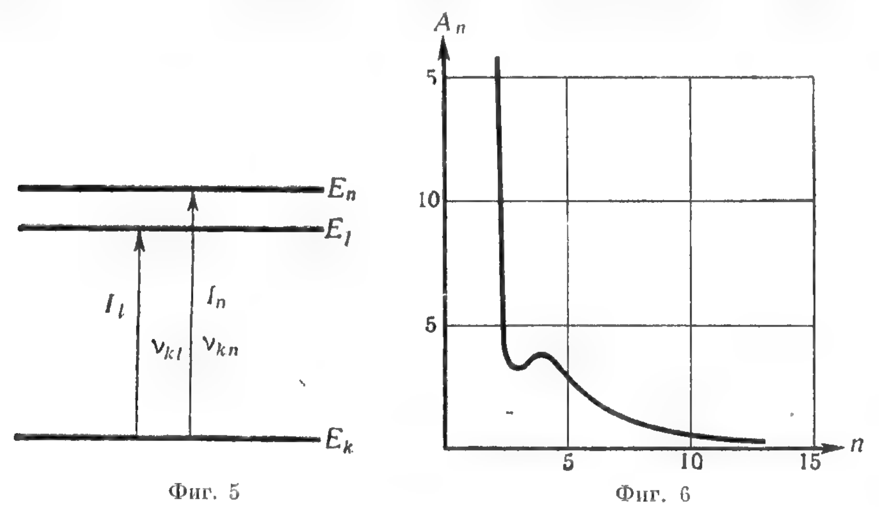
$$f_{kl} = \frac{N_l}{N_k},$$

где N_l находится, как было сказано, из промера положения «крюков». Для нахождения f_{kl} остается еще определить число поглощающих атомов N_k в единице объема. Это можно сделать, если проводить опыт с парами в виде однородного столба и знать их упругость. Однако во многих случаях достаточно знать относительные значения вероятностей переходов f_{kl}, тогда нет надобности определять N_k, и все измерения сводятся к нахождению места расположения «крюков».

Пусть две спектральные линии поглощения с частотами ν_{kn} и ν_{kl} возникают при переходе между одним нижним и двумя верхними уровнями (фиг. 5). Отношение их интенсивностей, по формуле (4*a*), равно

$$\frac{I_l}{I_n} = \frac{f_{kl}}{f_{kn}}, \qquad (6)$$

т. е. оно определяется лишь отношением вероятностей переходов f_{kl}/f_{kn}. Этим еще раз подтверждается, что метод Рождественского позволяет особенно просто определять отношения интенсивностей линий.



Фиг. 5 Фиг. 6

Для случая составляющих одного и того же спектрального дублета $A_{kl} = A_{kn}$; кроме того, если дублет узкий, то частоты ν_{kl} и ν_{kn} можно приближенно положить равными друг другу; тогда по (6) и (5)

$$\frac{I_l}{I_n} = \frac{g_l}{g_n}.$$

Величины g_l и g_n имеют в теории Бора смысл статистических весов энергетических уровней E_l и E_n и равны $2J+1$, где J — «внутреннее» квантовое число данного уровня. Для уровней, соответствующих дублетам щелочных металлов, J равно $\frac{3}{2}$ и $\frac{1}{2}$, откуда

$$\frac{I_l}{I_n} = \frac{g_l}{g_n} = \frac{2J_l+1}{2J_n+1} = \frac{4}{2} = 2.$$

Таким образом, теория Бора объясняет открытый Рождественским целочисленный характер отношения интенсивностей составляющих спектральных дублетов щелочных металлов.

Механизм испускания спектральных линий определяется двумя характерными для атома величинами: энергиями уровней E и вероятностями переходов f. Первые находятся на основании спектральной систематики, вторые — наиболее легко и точно методом «крюков» Рождественского. Отсюда становится ясным значение работ Рождественского для всего дальнейшего развития учения об атомах.

Одновременно с успехами теории Бора вскрылись и ее крупные недостатки. Объясняя целочисленные отношения интенсивностей линий, она — с ее представлениями о неизменных стационарных состояниях — не могла объяснить самого факта аномальной дисперсии. Кроме того, дальнейшие измерения Рождественского показали, что отношения интенсивностей не всегда целочисленны: для второго дублета рубидия получается отношение 2.6, а для третьего дублета цезия 7.4. Отступление от целых чисел здесь несомненно больше ошибок наблюдений. Таким образом, богатый опытный материал, даваемый работами Рождественского, неизменно переплетался со всеми деталями развития атомных теорий.

Выход из затруднения был дан квантовой механикой Шредингера — Гейзенберга, которая снова привела к формуле Зельмейера и показала, что возможны отступления от целочисленности отношений интенсивностей составляющих спектральных мультиплетов. Вместе с тем квантовая механика выдвинула требование систематического определения для атомов вероятностей переходов f, как одних из основных характеристик их свойств. Отсюда стал ясным и дальнейший путь развития экспериментальных работ по аномальной дисперсии: используя метод Рождественского, определить вероятности переходов для возможно большего числа различных атомов, а для каждого атома — для возможно большего числа линий. Эту задачу в значительной мере выполнили ученики и сотрудники Рождественского.

Из выполненных учениками и сотрудниками Рождественского работ необходимо особо отметить работы В. К. Прокофьева и А. Н. Филиппова по определению вероятностей переходов в ряде последовательных линий в главных сериях атомов щелочных металлов. Отношения интенсивно-

стей составляющих мультиплета с одним и тем же значением главного квантового числа, как было указано, представляют собой, вообще говоря, целые числа. Отношения интенсивностей с разными главными квантовыми числами подчиняются иному закону. Этот закон может быть найден теоретически в общем виде лишь для простейших атомных систем: водорода и сходных с ним ионов. Здесь, как показывает квантовая механика, интенсивность в пределах одной серии меняется обратно пропорционально кубу главных квантовых чисел. Для более сложных атомов возможны отступления от закона кубов. Знание этих отступлений важно для понимания характера взаимодействия между электронами в оболочках атома.

Однако определение интенсивностей в данных сериях представляет исключительно большие трудности. Метод фотографической фотометрии для таких измерений непригоден, так как он потребовал бы точного знания распределения чувствительности фотопластинки в широкой спектральной области. Метод «крюков» Рождественского и здесь оказывается самым простым и точным.

А. Н. Филиппов, безвременно скончавшийся талантливый и тонкий экспериментатор, построил интерферометр с флюоритовой оптикой. Это — уникальный прибор, позволяющий работать во всей спектральной области, вплоть до крайнего ультрафиолета. Пользуясь таким интерферометром, Филиппов и Прокофьев наблюдали аномальную дисперсию у 25 последующих членов главной серии натрия и измерили вероятности переходов у 16 из них. Закон обратной пропорциональности интенсивностей кубам главных квантовых чисел подтвердился, начиная с 11-го члена серии. Для первых членов интенсивности спадают гораздо быстрее. На головной дублет главной серии натрия приходится 98.2% от общей интенсивности, на все остальные линии — менее 2%. Теоретический расчет, выполненный Прокофьевым методом квантовой механики, дал для первых четырех членов главной серии прекрасное совпадение с экспериментальными результатами.

Далее Филиппов измерил вероятности переходов в 12 последующих членах главной серии лития. Результаты его измерений представлены на фиг. 6. Как видно, у лития наблюдается своеобразный ход вероятностей переходов: он не монотонен, вероятность перехода для третьего члена серии больше, чем для второго. Этот экспериментальный результат также нашел подтверждение в кванто-механических расчетах, выпол

ненных В. А. Фоком и М. И. Петрашень. Наконец, Филиппов и Прокофьев исследовали ход вероятностей переходов в побочных сериях таллия. У таллия побочные серии являются сериями поглощения, и в этом отношении они аналогичны главным сериям щелочных металлов. Для второй побочной серии было найдено монотонное спадание интенсивности; для первой побочной серии вероятность перехода оказалась для четвертого члена в 1.8 раза больше, чем для третьего. В конце обеих серий снова выполняется закон кубов.

Измерения Филиппова и Прокофьева позволили еще в одном отношении сопоставить выводы теории и опыта. По так называемому правилу сумм Томаса — Куна алгебраическая сумма всех вероятностей переходов с данного уровня должна равняться $\frac{1}{3}$ числа степеней свободы k излучающего атома:

$$\sum_{1}^{\infty} f_i = \frac{1}{3} k.$$

Филиппов и Прокофьев нашли, что для натрия, и особенно для таллия, это соотношение не оправдывается. Впоследствии этот результат был еще более точно подтвержден Г. С. Кватером. В. А. Фок показал, что такое расхождение между опытными данными и правилом Томаса — Куна вполне закономерно и что «правило» сумм требует существенного исправления.

Весьма интересное исследование по аномальной дисперсии было выполнено Г. С. Кватером. Атом таллия характеризуется тем, что у него на небольшом расстоянии над нормальным уровнем $^2P_{1/2}$ расположен метастабильный уровень $^2P_{3/2}$. При высокой температуре часть атомов окажется переведенной в метастабильное состояние $^2P_{3/2}$, и тогда в спектре поглощения обнаружатся две побочные серии. Чем выше температура, тем большее число атомов будет находиться в метастабильном состоянии $^2P_{3/2}$. Отношение числа атомов в состояниях $^2P_{1/2}$ и $^2P_{3/2}$ определяется законом Больцмана, являющимся одним из основных законов современной статистической физики. По закону Больцмана, числа дисперсионных центров N_1 и N_2, соответствующих обоим состояниям атомов таллия, относятся как

$$\frac{N_1}{N_2} = Be^{\frac{ch\Delta v}{kT}}, \tag{7}$$

где B — некоторая константа, Δv— разность термов $^2P_{1/2}$ и $^2P_{3/2}$, h — постоянная Планка, k — постоянная Больцмана, c — скорость света, T — абсолютная температура паров таллия. Метод «крюков», дающий возможность определить при разных температурах отношение числа дисперсионных центров N_1/N_2, тем самым позволяет проверить формулу Больцмана и найти отношение h/k.

Логарифмируя равенство (7), получим

$$\ln\frac{N_1}{N_2} = \ln B + c\,\frac{h}{k}\,\Delta v\,\frac{1}{T},$$

откуда видно, что логарифм отношения N_1/N_2, по Больцману, линейно зависит от $1/T$, причем угловой коэффициент зависит от значения h/k. Г. С. Кватер измерил аномальную дисперсию в слое паров таллия определенной упругости в интервале температур 1165—1380° К и получил ожидаемую зависимость $\ln\frac{N_1}{N_2}$ от $1/T$, причем он нашел для отношения h/k значение $4.82 \cdot 10^{-11}$ град. сек., что в пределах ошибки наблюдения совпадает с точным значением этого отношения, равным $4.787 \cdot 10^{-11}$ град. сек. До сих пор было известно несколько более или менее косвенных способов проверки формулы Больцмана, метод же «крюков» Рождественского дал наиболее прямой и точный способ проверки этого основного закона физики.

В последние годы Рождественский предпринял и выполнил новое большое исследование по аномальной дисперсии. Употреблявшаяся ранее экспериментальная методика позволяла исследовать лишь сравнительно легкоплавкие элементы. Вещество вводилось в эвакуированную кварцевую трубку и нагревалось электрической печью из нихрома или платины. Это позволяло достигать температур, не превышающих 1200° С. Была очевидна потребность перейти к более высоким температурам, включить в круг исследований трудно испаряемые элементы и поставить новые задачи по определению числа возбужденных атомов и степени ионизации. С этой целью Рождественский приступил к созданию новой

экспериментальной установки с вакуумной печью, типа печи Кинга, с телом накала в виде графитовой трубки.

Работа была связана с преодолением исключительных экспериментальных трудностей. В одну из ветвей интерферометра, тончайшего оптического прибора, требовалось ввести громоздкую печь, легко нарушающую строгий температурный режим, необходимый для получения устойчивой интерференционной картины.

После года напряженной работы был построен новый интерферометр с размером стоек в 30 см, т. е. в 5 раз больших, чем у прежних интерферометров. Такой большой интерферометр в высшей степени чувствителен к малейшей тряске, и для того, чтобы добиться хороших интерференционных полос, его пришлось смонтировать по двойной системе каменных плит, общим весом в полтонны, на резиновых прокладках. В одну из ветвей интерферометра вводилась печь диаметром около 40 см и длиной в 75 см. Печь питалась от 30-киловаттного вольтпонижающего трансформатора и при силе тока в 1800 ампер давала температуру в 3000° С. Такая высокая температура привела к новым экспериментальным затруднениям. Например, в местах стыка графитовой трубки с подводящими электродами при нагревании возникали деформации, приводившие к разрушению электродов или к порче контакта.

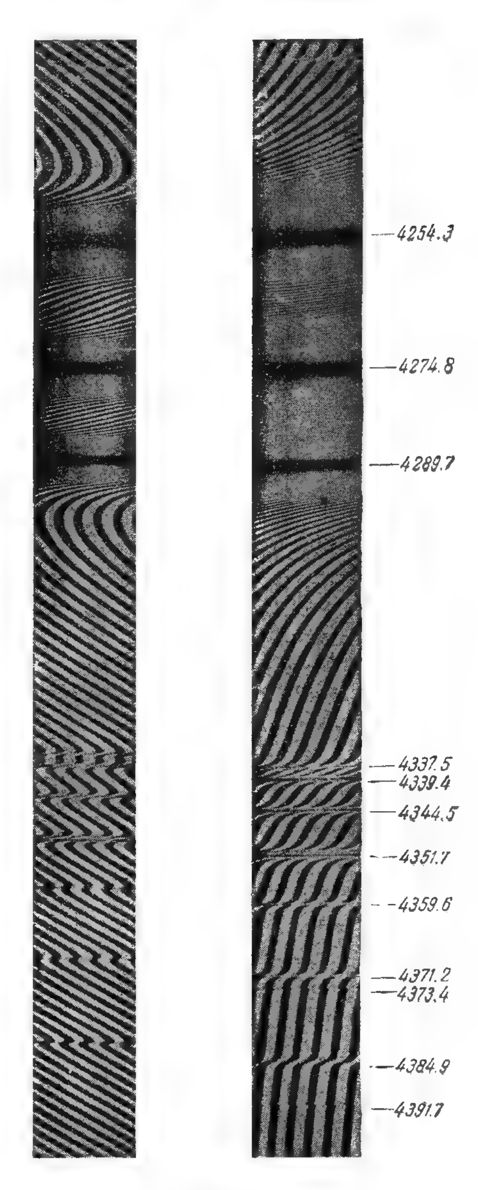
Трудности возникали также при создании оптической системы, обеспечивающей достаточную светосилу при наличии длинной и сравнительно узкой трубки. Тем не менее все они были преодолены, и новая установка позволила делать при высоких температурах (до 3000° С) снимки «крюков», вполне пригодные для промеров.

В результате проделанной работы экспериментальные возможности оказались сильно раздвинутыми, и Д. С. Рождественский наметил дальнейшую обширную программу исследований: прежде всего исследований аномальной дисперсии в парах элементов с достраивающейся d-оболочкой: Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu; затем наблюдений аномальной дисперсии на побочных сериях и на линиях ионизированных атомов. Совместно с Н. П. Пенкиным им было произведено качественное наблюдение аномальной дисперсии в парах железа и проведен ряд измерений в спектре хрома. Для весьма узкого мультиплета хрома $^5S_2 - {}^5P_{1,2,3}$ были измерены отношения интенсивностей, оказавшиеся равными 3.004 : 4.997 : 7.000, что совпадает с теоретическими целочисленными

отношениями 3 : 5 : 7. Кроме того, удалось качественно наблюдать аномальную дисперсию в побочных сериях калия и на линиях ионизованного бария. На фиг. 7 приведены фотографии интерференционных полос и «крюков» вблизи линий 4254.3, 4274.8, 4289.7 и др. хрома.

Н. П. Пенкин, продолжая наблюдения над парами хрома, измерил отношения интенсивностей составляющих мультиплета $^7S_3 - {}^7P_{2,3,4}$ и получил снова хорошее совпадение с теоретическими отношениями 5 : 7 : 9. Он исследовал вопрос о наличии в печи Кинга термодинамического равновесия. Кингом и другими авторами был проделан большой ряд работ с высокотемпературной печью, но при этом истинные условия в печи ни разу точно не выяснялись, что в значительной мере обесценивало результаты работ. Измерения по методу «крюков» показали, что вплоть до температуры в 2700° условия в печи близки к термодинамическому равновесию. Вновь была проверена выполнимость закона Больцмана и найдено значение для отношения h/k, близкое к установленному другими способами.

Наконец, в самое последнее время (1946) Н. П. Пенкин произвел интересные наблюдения над аномальной дисперсией у линий ионизованного бария. При введении в печь метал-



Фиг. 7

лического бария получилось хорошее согласие с теоретической формулой Саха — Фоулера; при введении солей бария (хлористый и азотнокислый) степень ионизации оказалась выше теоретической.

Таким образом, метод Рождественского оказался пригодным для решения новых задач, имеющих большое значение для ряда важнейших разделов физики и астрофизики.

Д. С. Рождественский открыл новый способ исследования аномальной дисперсии. Его школа является в СССР и во всем мире представительницей оригинального прецизионного направления спектроскопии. Его ученики в настоящее время работают над измерением атомных констант, знание которых необходимо для расширения наших сведений о строении материи.

ЛИТЕРАТУРА

1. *Д. С. Рождественский.* К исследованию аномальной дисперсии в парах натрия. ЖРФХО, Физ. отд., т. 42, 1910, стр. 87.
2. *Д. С. Рождественский.* Аномальная дисперсия в парах натрия. Магистерская диссертация. СПб., 1912.
3. *Д. С. Рождественский.* Простые соотношения в спектрах щелочных металлов. Докторская диссертация. Пг., 1915; Тр. Гос. Оптич. ин-та, 1921, т. 2, № 13.
4. *Д. С. Рождественский* и *В. И. Туроверов.* Соотношение двух линий D_1 и D_2 дублета натрия. Тр. Гос. Оптич. ин-та, 1921, т. 2, № 14.
5. *Д. С. Рождественский.* Аномальная дисперсия в широких полосах поглощения. Изв. АН СССР, серия физ., 1934, № 1, стр. 35.
6. *Г. С. Кватер.* Докторская диссертация. 1945.
7. *Д. С. Рождественский.* Анализ спектров и спектральный анализ. Изв. АН СССР, серия физ., 1936, № 1—2, стр. 189.
8. *В. К. Прокофьев, А. Н. Филиппов.* Zs. f. Phys., Bd. 56, 1929, S. 458.
9. *А. Н. Филиппов.* Тр. Гос. Оптич. ин-та, 1932, т. 8, № 85.
10. *В. А. Фок, М. И. Петрашень.* ЖЭТФ, 1934, т. 4, стр. 295.
11. *В. К. Прокофьев, А. Н. Филиппов.* Zs. f. Phys., Bd. 85, 1933, S. 647.
12. *Г. С. Кватер.* ЖЭТФ, 1935, т. 5, стр. 426.
13. *В. А. Фок.* Zs. f. Phys., Bd. 89, 1934, S. 744.
14. *Д. С. Рождественский, Н. П. Пенкин.* Определение сил вибраторов в спектрах атомов. Изв. АН СССР, серия физ., 1941, № 5, стр. 97.
15. *Н. П. Пенкин.* Изв. АН СССР, серия физ., 1946.

Академик

С. И. Вавилов

ЛЮМИНЕСЦЕНЦИЯ И ЕЕ ДЛИТЕЛЬНОСТЬ

§ 1

РАВНОВЕСНОЕ И НЕРАВНОВЕСНОЕ ИЗЛУЧЕНИЕ И ЛЮМИНЕСЦЕНЦИЯ


аблюдаемый свет всегда возникает в сложных статистических процессах, глубоко отличных от абстрактного теоретического образа изолированной элементарной системы, электрона, атома или молекулы. Имеется возможность регистрировать действие немногих, иногда отдельных фотонов даже в видимой части спектра [1], но нет способов управлять излучением одной частицы и наблюдать повторные акты радиации именно от нее. Принципиально доступны измерению частота и состояние поляризации отдельного излученного фотона, но исключена возможность исследования на одной и той же частице полного спектра, «длительности» свечения и пространственной направленности излучения, связанной с «мультипольностью» элементарного осциллятора. Все эти основные свойства излучения наблюдаемы только для очень сложного коллектива.

Такие реальные излучающие коллективы интересны и важны прежде всего именно своей реальностью и конкретностью. Кроме того, путь к пониманию свойств самих элементарных излучателей идет необходимо

через их излучение. Предельный и наиболее изученный случай светящегося коллектива — равновесное тепловое излучение черного тела — замечателен своей общностью, бесследно скрывающей все специальные черты элементарных излучающих частиц. Это излучение не дает никакой возможности экспериментального определения особенностей светящихся частиц, образующих черное тело.[1] Вместе с тем, с точки зрения наших теоретических представлений элементарные процессы, из которых слагается свечение черного тела, ничем не отличаются от элементарных процессов в других светящихся системах. Для доказательства достаточно «развернуть» черное тело или превратить равновесное свечение в неравновесное. Взаимные влияния элементарных процессов в условиях теплового равновесия таковы, что индивидуальные свойства частиц сливаются в единую характеристику коллектива.

Всякое тело, находящееся некоторое время в тепловом равновесии, характеризуемом определенной температурой, также служит источником равновесной, «не черной» радиации. На основании термодинамических соображений в соответствии с теоремой Кирхгофа «не черное» излучение в данном спектральном интервале равняется произведению черного излучения при той же температуре на поглощательную способность тела в этой спектральной области. Таким образом, в «не черном» равновесном излучении проявляются, хотя и в малой степени, особые свойства вещества, из которого тело построено. Однако есть признак, объединяющий любое равновесное излучение, черное и «не черное». Он состоит *в полной независимости такого излучения от качества энергии, поддерживающей постоянство температуры тела.* К этому признаку мы вернемся в § 2.

---

[1] Несколько не ясным остается вопрос о том, может ли изменение характера элементарных излучателей (их мультипольности) повлиять на интерференционные свойства равновесного излучения, ускользающие от термодинамической нивелировки. Характер интерференционной картины при встрече в начале широко расходящихся когерентных пучков резко зависит от рода осцилляторов [2]. Например, для квадруполей интерференционная картина имеет вид, глубоко отличный от соответствующей картины для диполей. Будет ли черное тело, построенное из диполей и квадруполей, давать одинаковую картину при интерференции широко расходящихся пучков? Для опытов такого рода необходимо, чтобы излучающее тело хотя бы в одном измерении было тоньше длины световой волны. Но может ли такое тело обладать свойствами черного тела?

Наряду с равновесными излучающими коллективами существует бесконечное разнообразие неравновесных коллективов. Выше уже отмечалось, что всякое тело, находящееся при некоторой температуре выше абсолютного нуля, есть равновесный излучатель. Хотя при изменениях температуры равновесие в телах устанавливается медленно, однако промежуточные тепловые состояния, существующие даже только миллионные доли секунды (например, в конденсированной искре [3]), все же могут однозначно характеризоваться определенными температурами и соответствующим им тепловым излучением. Это — неизбежный фон всякого реального неравновесного излучения.

В 1888 г. Э. Видеман [4] ввел для неравновесных процессов излучения термин *люминесценция*. По Видеману признаком люминесценции служит превышение излучательной способности тела в данном спектральном интервале над излучательной способностью равновесного процесса. Это определение требует некоторых уточнений и вместе с тем ограничений для того, чтобы его можно было привести в соответствие с реально применяемыми в физике представлениями о люминесценции.

Качественно видемановское определение неравновесного излучения вполне правильно, но для его количественного выражения необходимы добавочные предположения. В соответствии с духом видемановского определения мы предполагаем *аддитивность* обоих процессов. Обозначим через $E_{\lambda T}$ общую излучаемую энергию тела при температуре[1] T в спектральной области λ, через $\varepsilon_{\lambda T}$ — энергию равновесного излучения того же тела при тех же условиях.

Тогда энергия неравновесного излучения (по Видеману — люминесценции)

$$I_{\lambda T} = E_{\lambda T} - \varepsilon_{\lambda T}. \qquad (1)$$

Подчеркнем при этом предположение, что $E_{\lambda T}$ всегда больше $\varepsilon_{\lambda T}$. Иначе говоря, фон равновесного излучения принимается всегда существующим и меньшим $E_{\lambda T}$.

---

[1] При переходе от равновесного состояния тела к неравновесному, например при освещении тела посторонним светом, возбуждающим вторичное свечение, температура тела может несколько повыситься. Под T разумеется установившаяся температура неравновесного излучающего тела.

Если не отклоняться от фактического содержания понятия люминесценции, которым пользуется современная физика, то было бы неправильным отождествлять широкое понятие неравновесного излучения, определяемое (1) с конкретным понятием люминесценции.

В отличие от равновесного излучения, неравновесная радиация всегда носит характерные черты излучающего тела или причины, вызвавшей радиацию, или того и другого вместе. Особенности отражаются в спектре, в интенсивности, в поляризации в пространственной анизотропии (связанной с мультипольностью элементарных излучателей [5]) и, наконец, в длительности свечения.

Можно выделить два крайних случая, особенно резко указывающих на принципиальную разнородность процессов, обнимаемых общим понятием неравновесного излучения. Известны многочисленные случаи неравновесного свечения, в широких пределах не зависящего в отношении спектрального состава от способа возбуждения. Например, свечение кристаллов, содержащих радикал уранила, при возбуждении любыми световыми лучами, электронами любых скоростей, лучами альфа, механическим ударом и другими способами остается по спектральному составу неизменным. Здесь, независимо от природы возбуждающего агента, проявляется излучение, зависящее от химической природы вещества.

Противоположный крайний случай — рассеяние и отражение света. Еще Ньютон на тех объектах, которые были ему доступны, установил, что при рассеянии и отражении монохроматического света на любом веществе его спектральный состав остается неизменным [6], т. е. в этих ньютоновских опытах неравновесное излучение характеризовало только возбуждающую причину, а не состав тела.

Приведенные примеры отмечают крайности. Бесконечное разнообразие случаев неравновесного излучения несет на себе отпечаток как свойств излучающего вещества, так и особенностей возбуждающей причины. Почти всегда спектр, выход, поляризация и другие свойства излучения зависят от качества возбуждения. С другой стороны, вопреки Ньютону, обнаружились факты изменения спектра и других свойств возбуждающего света при рассеянии (явление Комптона, комбинационное рассеяние).

Таким образом, несмотря на несомненную разнородность процессов неравновесного излучения, их нельзя разделить по признаку не-

зависимости от возбуждающей причины или, наоборот, от особенностей тела.

Существуют ли другие свойства, доступные наблюдению и измерению, по которым можно было бы рационально провести такое разделение? До недавнего прошлого таким признаком, практически трудно наблюдаемым, но принципиально доступным, могли считаться когерентные свойства рассеяния и отражения в отношении возбуждающего света. Когерентность отличает эти процессы от флуоресценции и многих иных неравновесных процессов. Однако открытие некогерентных видов рассеяния Комптоном, Раманом, Мандельштамом и Ландсбергом превращает такого рода разделение процессов неравновесного излучения в непоследовательное.

Искомое разделение может быть, однако, осуществлено естественным, рациональным и, по крайней мере на сегодняшний день, вполне последовательным путем, если обратиться к признаку *длительности свечения*, т. е. послесвечения тела после прекращения или резкого изменения возбуждающего агента [7].

Если между возбуждающим агентом (первичным светом, возбуждающими электронами и т. д.) и неравновесным излучением не вклиниваются промежуточные процессы, то излучение должно возникать и исчезать практически мгновенно (точнее, за время порядка периода светового колебания). Наоборот, при наличии промежуточных процессов (возбужденное лабильное или метастабильное состояние атомов и молекул, ионизация с последующей рекомбинацией и т. д.) неравновесное излучение не может прекращаться сразу после прекращения возбуждения и вместе с тем не сразу достигает своего стационарного значения. К первому классу явлений относятся все случаи рассеяния света, отражения, тормозной радиации, а также особого вида излучения, возникающего при распространении электронов в среде со скоростью, превышающей фазовую скорость света (радиация Черенкова [8]). Во второй класс попадают разнообразные виды флуоресценции, фосфоресценции и других процессов, практически и объединяемых понятием люминесценции.

Пользуясь представлениями классической физики, можно сказать, что при наличии послесвечения в системе имеются *собственные* колебания, при отсутствии же послесвечения мы имеем дело с *вынужденными* колебаниями. В соответствии с этим всю часть неравновесного излучения,

не обнаруживающую послесвечения, рационально назвать *вынужденным излучением.*

На первый взгляд может показаться, что возможен непрерывный переход от вынужденного излучения к люминесценции, от практически мгновенно затухающего свечения к длительному. Например, меняя длину волны света, проходящего через среду в области, близкой к резонансной линии поглощения, мы переходим постепенно от рассеяния, не обладающего длительностью, к резонансному длительному излучению. В действительности при таком переходе нет постепенного превращения рассеяния в резонансное излучение, но только постепенное нарастание амплитуды всегда сосуществующих собственных колебаний. Длительность этих собственных колебаний остается неизменной на любом расстоянии от области резонанса.

Таким образом, разделение неравновесного излучения на люминесценцию и вынужденное излучение вполне последовательно и строго. Обозначим энергию люминесценции в спектральной области λ при температуре T через $L_{\lambda T}$ и энергию вынужденного излучения при тех же условиях $e_{\lambda T}$. В таком случае в качестве определения люминесценции можно написать

$$L_{\lambda T}=E_{\lambda T}-\varepsilon_{\lambda T}-e_{\lambda T}. \qquad (2)$$

Сравнивая (1) со (2), мы видим, что

$$L_{\lambda T}=I_{\lambda T}-e_{\lambda T}. \qquad (3)$$

Таким образом, признак длительности свечения позволяет расчленить любой процесс излучения на три принципиально различных слагаемых — *равновесную радиацию, люминесценцию и вынужденное излучение.* Не приходится сомневаться, что элементарные процессы, составляющие равновесную компоненту, могут быть в свою очередь принципиально разделенными на люминесцентные и вынужденные. Однако такое разделение по признаку послесвечения практически невыполнимо для равновесного свечения вследствие инерции теплового возбуждения. Наоборот, в неравновесном излучении, вызываемом световыми или корпускулярными лучами, прерывание возбуждения и наблюдение послесвечения вполне осуществимо даже до времен порядка примерно 10^{-10} сек.

§ 2

ЗАВИСИМОСТЬ ВЫХОДА ЛЮМИНЕСЦЕНЦИИ ОТ ЭНЕРГИИ ЭЛЕМЕНТАРНОГО ВОЗБУЖДЕНИЯ И ЗАКОН СТОКСА

В то время как для равновесного излучения вследствие особенностей коллектива, как указывается ниже, квантовый характер поглощения света не существенен, для люминесценции квантовая природа поглощения определяет ее вторую основную особенность, в частности находящую себе выражение в правиле Стокса.

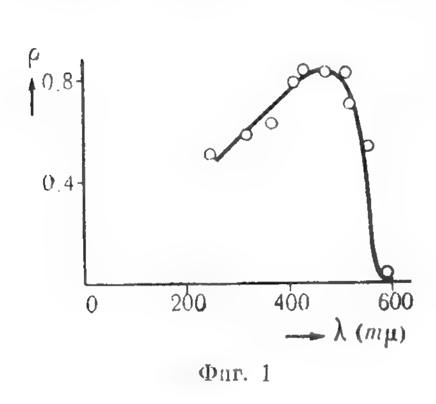
Качественное разнообразие видов возбуждения люминесценции (световое, корпускулярное, химическое) объединяется представлением об энергии элементарного акта возбуждения, которая может быть выражена во всех случаях соответствующим квантом $h\nu$. Имея это в виду, можно перейти к рассмотрению зависимости выхода люминесценции от энергии элементарного акта возбуждения.

Назовем выходом ρ_λ отношение излученной энергии к энергии, поглощенной в некотором спектральном интервале в области λ. Из самого определения равновесного излучения следует, что для него выход свечения не зависит от λ. Если тепловое равновесное излучение тела возмещается поглощением эквивалентной энергии, подводимой в виде радиации любой длины волны или корпускулярных лучей, то очевидно, что излучение в количественном и качественном отношении остается одним и тем же, независимо от качества поглощенной энергии. Иными словами, выход ρ не зависит от λ и равняется по определению 1. Это означает, что для любых видов поглощаемой энергии при одном и том же ее количестве физическое состояние тела в случае равновесия в среднем остается тем же. В связь с этим можно поставить то обстоятельство, что для вывода закона черного излучения не обязательно предположение о квантовом характере поглощения света, как в свое время показал М. Планк. Квантовые законы с необходимостью проявляются только в излучении черного тела.

Понятие о выходе свечения не имеет смысла для вынужденного излучения, так как в этом случае нет поглощения возбуждающей энергии, а происходит ее непосредственный, практически мгновенный переход

в энергию вынужденного излучения.[1] Наоборот, для люминесценции выход чрезвычайно характерен.

Зависимость выхода люминесценции от длины волны возбуждающего света экспериментально изучена до сих пор почти только для одного класса веществ, именно для растворов флуоресцирующих органических красителей. Впервые в достаточно широком спектральном интервале выход был исследован для растворов флуоресцеина [9]. Результаты изображены на фиг. 1. Мы видим там, что при переходе от коротких волн к более длинным выход растет, достигая максимума, и затем быстро падает практически до нуля. Позднее [10] были произведены аналогичные измерения для флуоресцеина, родамина, родулина, магдалы красной и эозина в различных растворителях и при разных температурах. При этом подтвердились результаты, найденные для флуоресцеина. Вместе с тем было показано, что крутизна падения выхода в области длинных волн быстро уменьшается при возрастании температуры. Наиболее обширные экспериментальные результаты по данному вопросу получены другим методом для большего числа органических веществ [11]. При этом найденный вначале экспериментальный закон изменения выхода от длины волны возбуждающего света подтвердился. Он оказался выполняющимся также для люминесценции уранового стекла. Таким образом, закон, изображаемый кривой на фиг. 1, повидимому, имеет универсальный характер, хотя следует заметить, что до сего времени нет измерений зависимости выхода от длины волны возбуждающего света для люминесценции газов и паров, а также кристалло-фосфоров.



Фиг. 1

Универсальность найденного закона может быть, однако, показана на основании весьма общих квантовых и термодинамических соображений [12].

[1] В случае комбинационного рассеяния, связанного с поглощением света, понятие выхода сохраняет значение, однако, совсем иного рода.

Примем, что возбуждение — монохроматическое с частотой ν', причем при поглощении света излучается непрерывный спектр с распределением энергии

$$\nu f(\nu)\,d\nu,$$

где $f(\nu)$, вообще говоря, зависит от ν'. Основываясь на квантовом характере поглощения и излучения света, можно утверждать (если только не происходит охлаждения среды при излучении), что

$$n'h\nu' \geqslant nh\int_0^\infty \nu f(\nu)\,d\nu,$$

где n' и n — соответствующие числа фотонов. Введем среднюю частоту излучения

$$\bar{\nu} = \frac{\int_0^\infty \nu f(\nu)\,d\nu}{\int_0^\infty f(\nu)\,d\nu},$$

отсюда $n'h\nu' \geqslant nh\bar{\nu}$ или

$$\rho = \frac{n\bar{\nu}}{n'\nu'} \leqslant 1.$$

Если

$$h\nu' > h\bar{\nu} \geqslant \frac{h\nu'}{2}, \tag{4}$$

то

$$n \leqslant n'.$$

Поэтому в пределах (4) в лучшем случае

$$n = n'.$$

Ограничимся этим случаем; тогда

$$\rho = \frac{\bar{\nu}}{\nu'}, \tag{5}$$

т. е. выход будет линейно и в равной мере возрастать, как $\bar{\nu}/\nu'$, в пределах выполнения (4). Такое линейное нарастание выхода должно, однако, нарушиться, как только мы переходим в антистоксову область, т. е.

$$\bar{\nu} > \nu'.$$

На основании принципов термодинамики можно доказать [12], что выход люминесценции не должен превышать единицы. Рассмотрим для доказательства, например, следующий мысленный опыт. Пусть в сосуде с идеально отражающими стенками заключено квазимонохроматическое излучение ν'. Часть сосуда покрыта слоями двух люминесцирующих веществ, которые можно, не совершая работы, закрывать отражающими крышками. Кроме того, эти слои можно изолировать от связанных с ними тепловых резервуаров также при помощи зеркал. Пусть одно из люминесцирующих веществ, отделенное на время данной операции от теплового резервуара зеркалом, поглощает свет, излучая антистоксову частоту ν с выходом большим единицы за счет самоохлаждения. Закроем после этого первый люминесцирующий слой зеркалом, восстановив его связь с тепловым резервуаром, и откроем второй слой, поглощающий свет с частотой ν и излучающий стоксов свет с прежней частотой ν' с выходом $\frac{\nu'}{\nu} < 1$. При этом, также на время операции, прекращаем связь с тепловым резервуаром при помощи зеркала. Избыток энергии нагревает второй люминесцирующий слой. Восстановим связь слоя с тепловым резервуаром и закроем слой снова зеркалом. Примем, что температура первого резервуара до опыта была ниже, чем второго. После опыта с первым люминесцирующим слоем и резервуаром температура во втором резервуаре еще повысится.

Таким образом, в случае осуществимости этих опытов было бы восстановлено первоначальное состояние в сосуде, и вместе с тем тепло без затраты работы было бы перенесено с менее теплого тела на более теплое в противоречие со вторым началом термодинамики. Причина ошибочного заключения, очевидно, кроется в предположении о возможности выхода, превышающего 1.

Изложенный мысленный опыт позволяет сделать и более детальные качественные выводы. Ясно, что если бы опыт давал положительные результаты, то чем меньше была возбуждающая частота ν', тем больше нарушалось бы второе начало, тем эффективнее происходило бы охлаждение холодного резервуара. Вместе с тем, чем ниже температура этого резервуара, тем глубже, т. е. ближе к абсолютному нулю, будет такое охлаждение. Отсюда естественно предполагать, что при уменьшении частоты ν' и понижении температуры при

опытах в антистоксовой области неравенство $\rho < 1$ должно усугубляться.

Этот вывод следует также из молекулярных соображений. Люминесцирующая среда, как изложено в § 1, характеризуется *сосуществованием* возбужденных состояний конечной длительности и теплового равновесия, соответствующего определенной температуре. Такому установившемуся состоянию в каждой возбужденной молекуле необходимо предшествует период *перераспределения* энергии. В это время существует коммуникация и обмен между электронной энергией возбуждения и тепловой энергией вращательно-колебательных степеней свободы молекулы. Структура молекулы определяет размер устанавливающейся энергии возбуждения $h\bar{\nu}$, как это можно в первом приближении представить в виде некоторой механической модели. Если поглощаемая энергия $h\nu' > h\bar{\nu}$ (стоксова область), то процесс ее обмена с тепловой энергией данной молекулы и среды протекает по естественному направлению: при устанавливающемся возбуждении часть поглощенной энергии переходит в тепло. В случае $h\nu' < h\bar{\nu}$ (антистоксова область) установление возбужденного состояния достижимо только за счет охлаждения молекулы и среды. Процессы такого рода возможны только в качестве флуктуаций, статистически компенсируемых обратными процессами полного перехода поглощенной энергии в тепло. Очевидно, что флуктуации тем менее вероятны, чем меньше частота ν' (в антистоксовой области) и чем ниже температура среды.

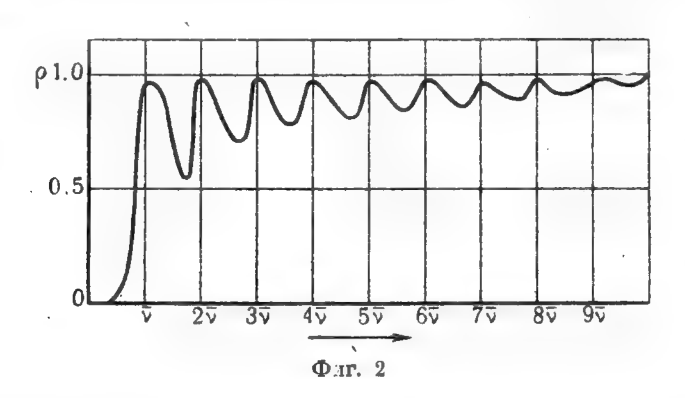
Мы получаем, таким образом, качественное объяснение экспериментально найденной зависимости (фиг. 1) и убеждаемся в ее универсальности.

До сих пор в стоксовой области возбуждения рассматривалась только часть спектра, удовлетворяющая (4). Перейдем теперь к большим частотам, когда

$$h\nu' \gg h\bar{\nu} \leqslant \frac{h\nu'}{2}. \qquad (6)$$

В этой части спектра возбуждения можно представить себе особые случаи, когда за счет поглощения $h\nu'$ образуются два и большее число излученных фотонов $h\bar{\nu}$. Посредством вторичного, третичного и т. д. излучения фотонов и электронов энергия большого (например, рентгеновского) кванта возбуждения в среде будет «размениваться», давая то

или иное число фотонов люминесценции $h\bar{\nu}$. С качественной стороны такой «размен» хорошо известен при возбуждении лучами Рентгена, катодными и радиоактивными излучениями. Однако количественно до сего времени эта область спектра возбуждения не изучена.

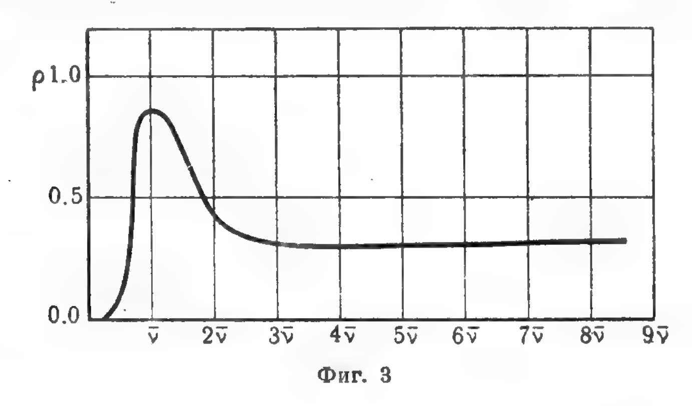


Фиг. 2

В идеальном случае, едва ли реализуемом, когда поглощаемая энергия полностью тратится на возбуждение и в тепло превращаются только квантовые остатки энергии, меньшие $h\bar{\nu}$, функция возбуждения должна приобретать сложный вид, представленный на фиг. 2. Не трудно заметить, что положения максимумов и минимумов на этой кривой должны приблизительно соответствовать $n\bar{\nu}$, где n проходит значения 1, 2, 3, 4, 5 и т. д. Величина минимальных значений ρ должна равняться $\frac{n-1}{n}$. В действительности поглощение и рассеяние вторичных радиаций в среде должно сглаживать функцию возбуждения в области больших частот и уменьшать абсолютное значение выхода ρ. Можно ожидать функции возбуждения вида, представленного на фиг. 3.

Необходимое и универсальное свойство люминесценции — это *резкий спад функции возбуждения со стороны длинных волн.* В связи с этим находится правило Стокса, установленное около века тому назад, но до сего времени остающееся «правилом», а не законом, т. е. утверждением не всегда справедливым и имеющим исключения. Правило Стокса в своем первоначальном виде гласит, что длина волны люминесценции всегда больше длины волны возбуждающего света. В таком виде это правило

нарушается очень часто, особенно если позаботиться о полном поглощении возбуждающего света и в антистоксовой области. Другая формулировка правила Стокса более гибка, она сводится к положению о том, что максимум полосы излучения смещен в сторону длинных волн по



Фиг. 3

отношению к максимуму спектра поглощения. Однако и в таком виде правило Стокса имеет исключения [13].

Как легко понять на основании изложенного выше, действительное содержание правила Стокса соответствует необходимости резкого спада функции возбуждения в антистоксовой области. Как только частота возбуждения ν' становится меньше средней частоты излучения $\bar{\nu}$, выход должен падать и тем больше, чем меньше ν'. Излучение с длиной волны меньшей, чем возбуждающаяся, возможно, но выход его будет тем меньше, чем меньше излучаемая средняя волна.

§ 3

СВОЙСТВА ФОТОЛЮМИНЕСЦЕНЦИИ РАСТВОРОВ И ДЛИТЕЛЬНОСТЬ ВОЗБУЖДЕННЫХ СОСТОЯНИЙ

Конечная длительность послесвечения, определяемая длительностью возбужденных состояний, служит, как показано в § 1, рациональным критерием выделения люминесценции из прочих процессов, составляющих неравновесное излучение.

Вместе с тем конечная длительность во многом определяет основные свойства люминесценции, доступные прямому наблюдению (спектр, выход, поляризацию). Интенсивность полос поглощения и излучения в спектрах люминесцирующих тел непосредственно связана с декрементом затухания элементарных осцилляторов в классической физике или с длительностью возбужденных состояний в современной квантовой интерпретации. Элементарное излучение разного типа (дипольное, квадрупольное и т. д.) соответствует резко отличающейся длительности возбужденных состояний. Течение во времени элементарных процессов, связанных с люминесценцией, определяет наиболее характерные черты люминесценции всех видов в газах и парах, в растворах, в кристаллах и полупроводниках.

С этой точки зрения в виде примера ниже излагаются на основании экспериментальных и теоретических работ последних десятилетий некоторые свойства фотолюминесценции растворов.

Поляризация флуоресценции растворов и ее зависимость от вязкости среды дает своей непосредственностью особенно наглядный пример связи свойств люминесценции с длительностью возбужденных состояний. Молекула, излучающая поляризованную флуоресценцию за время возбужденного состояния, поворачивается вследствие броуновского движения. Направление вращения различно у разных молекул, подчиняясь статистическим законам. Чем меньше вязкость среды, тем больше величина такого поворота; направленные поляризованные световые колебания развертываются в итоге веером, что приводит к постепенной деполяризации излучения при уменьшении вязкости среды. На основе теории броуновского движения эти представления дают возможность установить количественную связь между степенью поляризации флуоресценции p, средней длительностью возбужденных состояний τ и вязкостью η [14]. В простейшем виде эта связь выражается формулой

$$\frac{1}{p}=\frac{1}{p_0}+\left(\frac{1}{p_0}-\frac{1}{3}\right)\tau\frac{\varkappa\theta}{v\eta}, \qquad (7)$$

где p_0 — предельная поляризация при $\eta\to\infty$, определяемая на опыте, $\varkappa$ — постоянная Больцмана, θ — абсолютная температура, v — молекулярный объем. Формула (7) много раз проверялась на опыте [15] и

может считаться вполне подтвержденной. Если известны p_0 и η, то (7) дает удобный способ экспериментального определения τ_0.

Наряду с броуновским вращением поступательное броуновское движение молекул также дает способ измерения τ. Основой служит известное явление тушения флуоресценции растворов, т. е. уменьшение выхода флуоресценции при добавлении некоторых посторонних веществ (например, иодистого калия, анилина и др.).

С точки зрения теории, развитой в основной части в 1929 г. и в настоящее время, повидимому, общепринятой [16], тушение посторонними веществами соответствует взаимодействию возбужденных молекул с молекулами тушителя. Это взаимодействие может иметь характер физических «ударов второго рода» или очень кратковременных химических соединений. Применение теории диффузии и теории броуновского движения приводит к следующему выражению для выхода флуоресценции раствора в зависимости от концентрации тушащего вещества c:

$$\frac{\rho_0}{\rho} = (1 + w\tau_0 4\pi D\sigma Nc)\,\delta, \tag{8}$$

где ρ — выход фотолюминесценции в присутствии тушащего вещества, ρ_0 — выход в отсутствие тушителя, w — вероятность тушения при встрече возбужденной молекулы с молекулой тушителя, D — коэффициент диффузии, σ — радиус сферы действия возбужденной молекулы, N — число Авогадро, c — число граммов растворенного тушителя в 1 см³

$$\frac{1}{\delta} = 1 - 2e^{\gamma^2}\left(\frac{\sqrt{\pi}}{2} - \int_0^{\gamma} e^{-z^2}dz\right),$$

$$\gamma = \frac{w4\pi D\sigma^2 Nc}{2\sqrt{\pi D}\sqrt{\frac{1}{\tau_0} + w4\pi D\sigma Nc}},$$

$$D = \frac{\varkappa\theta}{12\pi\eta}\left(\frac{1}{\sigma} + \frac{1}{r}\right)$$

(r — радиус молекулы тушителя).

Формула (8) связывает выход люминесценции при наличии тушителя с τ_0 и подтверждается опытом во всяком случае для не слишком больших вязкостей η и концентраций c.

Теория поляризации свечения в зависимости от τ и теория тушения люминесценции посторонними телами, результаты которой содержатся в формуле (8), дают полное объяснение другого свойства флуоресценции растворов: возрастания поляризации по мере увеличения тушения люминесценции посторонними телами. Длительность возбужденного состояния, характеризуемая средней величиной τ_0, различна у разных молекул. С точки зрения изложенного представления о механизме тушения люминесценции посторонними веществами вероятность тушения возбужденных молекул с большим τ соответственно возрастает. Поэтому в результате тушения высвечиваться будут преимущественно частицы с малыми значениями τ. Но, согласно теории поляризации, уменьшение τ должно сопровождаться возрастанием поляризации, как это ясно из формулы (7). Для простейших случаев можно вывести общее соотношение [17]

$$\frac{\rho}{\rho_0}=\frac{\tau}{\tau_0}. \tag{9}$$

Подставляя его в (7), приходим к выражению, связывающему поляризацию, выход и τ_0

$$\frac{1}{p}=\frac{1}{p_0}+\left(\frac{1}{p_0}-\frac{1}{3}\right)\tau_0\frac{\rho}{\rho_0}\frac{\varkappa\theta}{v\eta}. \tag{10}$$

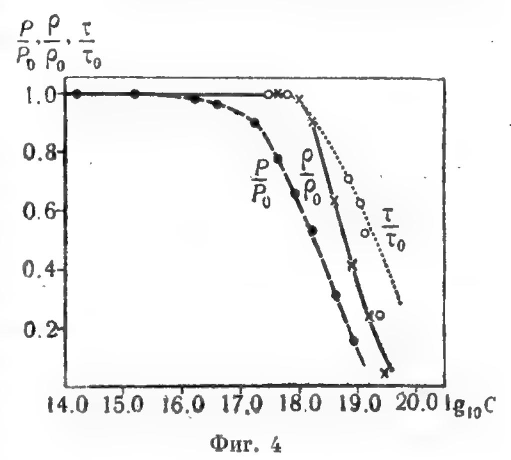
Формула (10) также подтверждается опытом и дает новый метод определения τ_0 [18].

В рассмотренных до сих пор явлениях, определяемых длительностью свечения, нет зависимости между молекулами люминесцирующего вещества. Излучающий неравновесный коллектив объединяется только общими статистическими законами и свойствами среды (вязкостью, температурой). В действительности во многих случаях флуоресценции растворов мы встречаемся с резким проявлением взаимодействия растворенных молекул, способных светиться.

Поглощение возбуждающего света во флуоресцирующем растворе, отнесенное к одному и тому же числу поглощающих молекул, остается обычно неизменным в очень широком интервале изменения концентраций (закон Бэра); между тем основные свойства излучения — выход, поляризация и длительность — резко изменяются при увеличении концентрации, начиная с очень малых концентраций. На фиг. 4 сопоставлены для раствора флуоресцеина в глицерине кривые изменения выхода ρ, степени

поляризации p и τ при изменении концентрации. По оси абсцисс отложены логарифмы концентраций, выраженных числом молекул в 1 см³. Приведенные кривые типичны для довольно широкого круга изученных до сих пор органических веществ, в особенности красителей.

Влияние концентрации на p, ρ и τ имеет необычный характер: сначала оно почти незаметно и только, начиная с определенной концентрации, резко возрастает. При этом раньше всего сближение молекул сказывается на поляризации.



Фиг. 4

Явление концентрационного тушения флуоресценции было известно еще Стоксу; концентрационная деполяризация и уменьшение τ обнаружены недавно [19]. До последнего времени делаются попытки свести влияние концентрации на свойства флуоресценции растворов к различным химическим и физико-химическим причинам (коагуляция, высаливание, электролитическая диссоциация и т. д.). Однако обратимость этих явлений (при восстановлении прежней концентрации) и другие количественные и качественные особенности, излагаемые ниже, исключают правильность таких толкований.

Один факт начала деполяризации свечения при чрезвычайно малых концентрациях, когда взаимодействующие молекулы удалены друг от друга более чем на 10—20 молекулярных радиусов, исключает правдоподобие физико-химических объяснений.

За последние годы была развита теория резонансной миграции энергии возбуждения во флуоресцирующих растворах. Представление о резонансных переходах возникло при исследовании резонансного излучения в парах и газах и предлагалось также на основе некоторых специальных предложений для растворов [20]. Конкретные предположения, сделанные в этом направлении, были, однако, ошибочными, и количественно правильных выводов получить не удалось.

Возможно, однако, построить общую теорию, без специальных гипотез, лишь на основе предположения о конечной вероятности переноса энергии, зависящей от расстояния и от времени взаимодействия, и о статистическом беспорядке явления [21].

Теория дает для вероятности $q\,(c,\ t)$ отсутствия резонансного переноса энергии от возбужденной молекулы другим невозбужденным за время t от начала возбуждения и при концентрации c следующее выражение:

$$q\,(c,\ t)=e^{-\lambda_0 c}e^{-\frac{ct}{k_2}}, \tag{11}$$

где λ_0 и k_2 — две постоянные, определяемые структурой молекулы и среды. Для вероятности $w\,(c,\ t)$ переноса энергии возбуждения, сопровождающегося тушением, теория приводит к аналогичному выражению

$$w\,(c,\ t)=e^{-\omega_0 c}\,e^{-\frac{ct}{k_1}}, \tag{12}$$

где ω_0 и k_1 — две другие постоянные.

Значения вероятностей q и w дают возможность при некоторых упрощающих допущениях (позволительных для не очень больших концентраций) найти выражения для зависимости степени поляризации, выхода и τ_0 от концентрации c и τ. Эти выражения (приближенно) имеют следующий вид:

$$\frac{1}{p}=\frac{1}{p_0}+\frac{c\tau_0}{k_2 p_0}\,(1-\pi), \tag{13}$$

где $\pi=\frac{p'}{p_0}$, причем через p' обозначено значение степени поляризации после первого резонансного перехода энергии возбуждения

$$\frac{\rho_0}{\rho}=1+c\left(\frac{\tau_0}{k_1}+\Omega_0\right)+c^2\left(\frac{\tau_0}{k_1}+\frac{\tau_0}{k_2}\right)\Omega_0+c^3\,\frac{\tau_0}{k_2}\,\Omega_0^2, \tag{14}$$

где

$$\Omega_0=\lambda_0+\omega_0,$$

$$\tau=\frac{\tau_0 k_1}{c\tau_0\left\{1+\frac{k_1}{k_2}(1-e^{-\omega_0 c})\right\}+k_1}. \tag{15}$$

Таким образом, поляризация, выход и средняя длительность при данной концентрации оказываются зависящими от τ_0. В руках экспериментатора

имеется при этом возможность менять τ_0 в довольно широких пределах посредством прибавления к раствору посторонних тушащих веществ, сокращающих τ_0 в соответствии с (11).

Формулы (13), (14), (15) содержат четыре эмпирические постоянные k_1, k_2, ω_0, λ_0. Надо иметь в виду, однако, что три формулы, в которые входят одни и те же постоянные, соответствуют трем независимым явлениям и измерениям. Поскольку во многих случаях $\lambda_0=0$, три эти измерения позволяют однозначно определить постоянные

$$\omega_0, \; k_1, \; k_2.$$

Систематические экспериментальные данные, полученные для многих веществ, подтверждают найденные соотношения [22]. Теория резонансной миграции энергии возбуждения в растворах подтверждается вместе с тем и другими фактами, имеющими убедительность experimentum crucis.

Возможны и действительно предполагались три принципиально отличные гипотезы для объяснения влияний концентрации на свойства флуоресценции:

1) физико-химические процессы, протекающие еще до освещения и происходящие между невозбужденными молекулами; таковыми могут быть ассоциативные процессы всякого рода, электролитическая диссоциация;

2) столкновения и встречи между возбужденной и невозбужденными молекулами, приводящие к процессам, аналогичным «ударам второго рода»;

3) резонансные переходы энергии возбуждения от возбужденной молекулы к невозбужденным.

Предположим, что флуоресцирующий раствор налит в стеклянные, чрезвычайно тонкие капилляры, с радиусами $< 10^{-5}$ см, причем капилляры отделены друг от друга стеклянными стенками на значительно бóльшие расстояния. Системой капилляров такого рода могут служить пористые стекла, получаемые специальной химической обработкой.[1] Флуоресцирующий раствор, налитый в капилляры, находится в особом «анизотропном» состоянии. При достаточно малой концентрации флуоресцирующие молекулы располагаются здесь одна за другой только в *одном измерении*. Анизотропия такого рода не может иметь никакого значения

[1] Такие стекла получены и исследуются в лаборатории И. В. Гребенщикова в Гос. Оптическом институте.

для свойств флуоресценции, если влияние концентрации сводится к первому фактору, действующему еще до введения раствора в капилляры.

Наоборот, одномерность размещения молекул чрезвычайно существенна для двух других возможных причин концентрационного влияния. Количество диффундирующего вещества к возбужденной молекуле, если оно возможно только в одном измерении, будет значительно меньше, чем при изотропном пространственном распределении. Точно так же резонансные влияния, ограничиваемые одним направлением, должны резко уменьшиться. Опыты такого рода [23] показали, что в пористых стеклах явления деполяризации и тушения флуоресценции растворов резко уменьшаются, что сказывается в очень значительном сдвиге критической концентрации, с которой явления начинают обнаруживаться в резкой форме, в сторону больших концентраций.

Такой результат мог бы толковаться как подтверждение второй гипотезы, сводящей концентрационные явления к столкновению молекул. Однако подсчет на основании теории диффузии совершенно исключает возможность такого объяснения с количественной стороны. Таким образом, допустимой остается только третья, резонансная гипотеза.

Другим важным следствием резонансной теории, выполнение которого может рассматриваться как решающий аргумент в ее пользу, служит необходимость деполяризации люминесценции при затухании [24]. Чем дальше стадия кривой затухания свечения, соответствующая излучению данной молекулы, тем в среднем бо́льшее число резонансных переходов энергии возбуждения предшествовало излучению. Если молекулы в среде расположены беспорядочно, то каждый переход несколько меняет направление поляризации, что в итоге выражается постепенной деполяризацией излучения.

Следует заметить, что в маловязкой среде еще бо́льшая деполяризация по мере затухания должна вызываться броуновским вращением возбужденных молекул [25]. Поэтому опыт такого рода может быть осуществлен только в очень вязкой среде, где броуновское вращение практически исключено. Деполяризацию при затухании впервые удалось наблюдать в некоторых сортах урановых стекол [26], обнаруживающих довольно сильно поляризованную люминесценцию. Урановое стекло — особенно удобный объект для наблюдения деполяризации при затухании вследствие того, что среднее τ в данном случае относительно велико

(порядка 10^{-4} сек.) и затухание может наблюдаться при помощи беккерелевского фосфороскопа. С другой стороны, вязкость стекла при комнатной температуре огромна. На опыте при затухании свечения урановых стекол степень поляризации убывала постепенно с 22 до 5% за две тысячных секунды. Существование деполяризации при затухании и зависимость ее от концентрации недавно удалось доказать и для растворов органических красителей [32] при помощи флуорометра, дающего возможность измерять длительности послесвечения до 10^{-10} сек. Эти опыты вместе с описанными или упомянутыми ранее доказывают с несомненностью резонансную природу концентрационных влияний в явлениях флуоресценции растворов.

Резонансная миграция энергии возбуждения во флуоресцирующих растворах служит характерным выражением особенностей излучающего коллектива, резко отличающим свойства системы от свойств изолированной молекулы. Энергия, поглощаемая одной молекулой, излучается другой, находящейся на большем расстоянии; поляризация оказывается в коллективе существенно иной. При этом основное значение для проявления этих особенностей коллектива имеет длительность возбужденных состояний.

§ 4

ЭВРИСТИЧЕСКОЕ ЗНАЧЕНИЕ КРИТЕРИЯ ДЛИТЕЛЬНОСТИ КАК ОСНОВНОГО СВОЙСТВА ЛЮМИНЕСЦЕНЦИИ

Конечное послесвечение как необходимое свойство люминесценции существенно не только как признак для рациональной классификации излучающих коллективов и как основа для теоретического объяснения многих особенностей люминесцирующих систем; оно может иметь решающее практическое значение для быстрого принципиального выяснения природы различных видов свечения. Условия для этого определяются практической доступностью критерия длительности в виде прямых методов (флуорометров и фосфороскопов) или косвенных; некоторые косвенные методы затронуты в § 3.

Такая роль критерия длительности при изучении различных видов свечения лучше всего пояснится двумя примерами, имеющими принципиальную важность.

В 1928 г., вскоре после открытия комбинационного рассеяния света, было опубликовано описание синего свечения с непрерывным широким спектром, возникавшего при распространении света в глицерине и амиловом алкоголе, многократно дестиллированных в вакууме [27]. Это свечение было истолковано как новый тип рассеяния света, т. е. относилось по терминологии § 1 к классу «вынужденного свечения». Как будто бы в подтверждение этой гипотезы «синее свечение» оказалось универсальным. Оно наблюдается почти во всех «чистых» жидкостях, в воде, спиртах, различных органических веществах, в серной кислоте и т. д., обладая всюду почти одинаковым спектром [28]. Критерий длительности позволил, однако, быстро опровергнуть сделанное предположение. Свечение можно тушить посторонними веществами, в частности иодистым калием, причем тушение подчиняется формуле (8), выведенной для тушения возбужденных состояний, обладающих конечной длительностью. Средняя длительность τ_0, определяемая по измерениям тушения, оказывается величиной порядка 10^{-9} сек., т. е. соответствует обычной молекулярной флуоресценции [28]. Далее удалось установить, что эта флуоресценция вызывается растворением различных органических остатков, носящихся в воздухе. «Универсальность» свечения объяснилась однотипным характером этих органических остатков. Критерий длительности перенес загадочное свечение из области гипотетического нового типа рассеяния в область привычной флуоресценции растворенных органических веществ.

Другой пример — прямо противоположного характера. Еще в конце XIX века, вскоре после получения радия, было замечено, что излучение солей радия вызывает в окружающих прозрачных средах заметное, хотя и очень слабое видимое свечение. Светятся, в частности, сами кристаллы солей радия и их водные растворы. Это свечение хорошо было известно супругам Кюри, открывшим радий. Кюри, как и другие исследователи, до последнего времени считали это свечение люминесценцией, вызываемой излучением радия.

Свечение такого рода было обнаружено в 1933 г. в чистой серной кислоте, помещенной в платиновый тигель и облучаемой лучами гамма-радия. Оно было найдено затем во многих прозрачных жидкостях, причем всюду оказалось, что интенсивность свечения в разных веществах при тех же прочих условиях мало изменяется [29]. Эти опыты производились через несколько лет после того, как была установлена универсальность

«синего свечения» жидкостей, вызываемого органическими остатками из воздуха. Поэтому естественным сначала было предположение о том, что свечение под действием лучей гамма тождественно с указанным «синим свечением», т. е. вызывается флуоресценцией растворенных органических молекул. Для решения вопроса был привлечен, так же как и в случае «синего свечения», критерий длительности.

Как говорилось выше, «синее свечение» в растворах тушится некоторыми посторонними веществами, например, иодистым калием, обнаруживая этим длительность возбужденных состояний. Опыты с новым свечением показали в отличие от этого, что оно не тушится ни при каких условиях (применение различных тушителей, уменьшение вязкости, повышение температуры).

Таким образом, было доказано, что новая радиация совершенно отлична от синего свечения и что она не обладает измеримой длительностью [29, 30] (длительность во всяком случае меньше 10^{-11} сек.). При сравнении свойств свечения, вызываемого лучами гамма, с «синей флуоресценцией» был найден факт поляризации нового свечения, причем плоскость поляризации оказалась повернутой на прямой угол по сравнению с поляризацией молекулярной флуоресценции в обычных условиях. Самый факт поляризации дал возможность нового применения критерия длительности. Если свечению молекулы предшествует возбужденное состояние конечной длительности, то при уменьшении вязкости раствора вследствие броуновского вращения молекул степень поляризации необходимо должна понижаться, если только средняя длительность τ_0 остается постоянной (§ 3). Измерения показали, однако, что при нагревании жидкости (например, серной кислоты и глицерина) поляризация нового свечения остается неизменной.[1] Этот опыт еще раз подтвердил, что исследуемое свечение не обладает измеренной длительностью, т. е. не может быть люминесценцией. *Свечение имеет вынужденный характер.*

---

[1] При нагревании «чистых» жидкостей, обнаруживающих «синее свечение», поляризация также изменяется мало. Но при этом происходит резкое понижение яркости вследствие увеличивающегося тушения, и τ_0 соответственно уменьшается. Такое сокращение τ_0 компенсирует возрастающее при повышении температуры броуновское вращение, вследствие чего поляризация мало изменяется. При возбуждении лучами гамма и повышении температуры жидкости яркость видимого свечения остается неизменной.

Таков был вывод применения критерия длительности к новой радиации. Явление, считавшееся в течение нескольких десятилетий почти тривиальным частным случаем люминесценции под действием жестких лучей, неожиданно предстало перед исследователями как нечто совершенно новое. Сначала была сделана попытка толковать свечение как результат торможения электронов рассеяния (комптоновских электронов в жидкости [30]). Эта гипотеза, объясняя некоторые стороны явления качественно, расходилась с опытом с количественной стороны. Она имела, однако, большое значение для дальнейшего исследования, так как перенесла первичное возбуждение радиации с лучей гамма на электроны. Отсюда непосредственно следовали опыты с влиянием магнитного поля (отклоняющего электроны) на характер свечения. Эти опыты привели к открытию, пожалуй, наиболее замечательного свойства новой радиации. Она оказалась направленной и сосредоточенной в пределах полого конуса с осью, совпадающей с движением возбуждающих электронов и повернутого вперед [8].

Изложенные экспериментальные этапы работы подготовили многое для создания теории. Несмотря на своеобразие и сложность явления, оно оказалось в своих основных чертах «классическим», т. е. объяснимым на почве классической теории электронов и учения об электромагнитном поле [31]. В основе теории лежит предположение, что *движущийся в среде* электрон может излучать даже при равномерном движении, если его скорость превосходит фазовую скорость распространения света в среде. В начальной стадии развития теории это заключение казалось гипотезой, затем его удалось вывести из классических представлений. Все свойства наблюденного излучения получили в теории полное количественное объяснение.

Можно провести довольно тесную аналогию между описанным «свечением Черенкова» и известной «головной волной» в акустике, возникающей в том случае, если из артиллерийского орудия вылетает снаряд со сверхзвуковой скоростью.

Лорд Кельвин, опираясь на артиллерийские опыты со звуком, до известной степени предвидел существование оптического аналога «головной волне». В знаменитой речи, сказанной в самом начале нашего века, в 1900 г., Кельвин говорил: «Если равномерная скорость атома превзойдет хотя бы немного скорость света, то должна появиться непериодиче-

ская, коническая волна... Угол конуса в эфире, так же как и для звука в воздухе, должен быть равным углу, синус которого есть отношение волновой скорости к скорости движущегося тела». Кельвин ошибся, полагая, что такое явление возможно в «эфире», т. е. в пространстве, лишенном вещества. В таких условиях скорость электрона не может превзойти фазовую скорость света. Такую же ошибку несколькими годами позже повторил Зоммерфельд. Но ожидания Кельвина реализовались в вещественной среде, где скорость электрона может превосходить фазовую скорость света.

Новая радиация открыла непредусматривавшуюся ранее страницу оптики «сверхсветовых скоростей». Вместе с тем новое явление постепенно, по мере его понимания и освоения, начинает получать существенное значение при исследовании космических лучей, а также и в других областях физики.

Рассмотренные примеры достаточно поясняют практическую эффективность критерия длительности для решения вопросов, связанных с природой различных видов излучения. Только этот критерий позволил перевести видимое свечение, возбуждаемое лучами гамма, из разряда тривиальных случаев люминесценции на уровень принципиально новых явлений. Есть основания надеяться, что он окажется полезным и решающим и в дальнейшем развитии физики и оптики.

ЛИТЕРАТУРА

1. *Е. М. Брумберг* и *С. И. Вавилов.* Изв. АН СССР, Отд. мат. и ест. наук, 1933, стр. 919; ДАН СССР, т. **1**, 1934, стр. 4; *С. И. Вавилов.* Тр. конфер. по физиол. оптике, 1934; Изв. АН СССР, серия физ., 1936, № 1/2, стр. 176; *С. И. Вавилов.* ДАН СССР, т. **21**, 1938, № 8; *S. Hecht, S. Schlaer* a. *M. H. Pirenne.* Science, June 20, 1941, p. 585; *Е. М. Брумберг, С. И. Вавилов* и *З. М. Свердлов.* ЖЭТФ, 1942, т. **12**, стр. 93; *С. И. Вавилов* и *Т. В. Тимофеева.* ЖЭТФ, там же, стр. 105 и 109.
2. *С. И. Вавилов.* Изв. АН СССР, ОМЕН, 1932, стр. 1451; *S. I. Wawilow* u. *E. M. Brumberg.* Sow. Phys., 1933, vol. 3, p. 103; *С. И. Вавилов.* ДАН СССР, т. **17**, 1937, стр. 159; *Q. Halpern* a. *F. W. Doerman.* Phys. Rev., vol. **52**, 1937, p. 937; *F. W. Doermann.* Phys. Rev., vol. **55**, 1939, p. 486; *S. I. Weissmann.* Phys. Rev., vol. **60**, 1941, p. 440.
3. Ср., например, *Н. Н. Соболев.* ЖЭТФ, 1943, т. **13**, стр. 137.
4. *E. Wiedemann.* Wied. Ann. d. Phys., 1888, vol. **34**, p. 446; 1889, vol. **37**, p. 177.
5. *С. И. Вавилов.* ЖЭТФ, 1940, т. **10**, стр. 1363; *П. П. Феофилов.* ДАН СССР, т. **55**, 1947, стр. 407.

6. *С. И. Вавилов.* Изв. АН СССР, серия физ., 1943, т. 7, стр. 2.
7. *С. И. Вавилов.* Электричество, 1945, № 1/2.
8. *П. А. Черенков.* Тр. Физич. ин-та, 1944, т. 2, стр. 2.
9. *S. I. Wawilow.* Zs. f. Phys., Bd. 42, 1927, S. 311.
10. *Л. А. Тумерман.* Тр. ФИАН, 1938, т. 1, стр. 77.
11. *С. С. Соломин.* ДАН СССР, т. **31**, 1941, стр. 741.
12. *S. Vavilov.* Journ. of Phys., 1945, vol. **9**, p. 68; критические замечания на эту статью: *P. Prinojsheim.* Ib., 1946, vol. **10**, p. 495. Ответ на эти замечания: *S. Vavilov.* Ib., 1946, vol. **10**, p. 499. *L. Landau.* Ib., 1946, vol. **10**, p. 503.
13. *N. Prilesschajewa.* Acta Phys. Chim., 1935, vol. **1**, p. 785.
14. *В. Л. Левшин.* ЖРФХО, физ. секц., т. **57**, 1925, стр. 283; *S. I. Wawilow.* Zs. f. Phys., Bd. **31**, 1925, p. 764; *F. Perrin.* Journ. de Phys., 1926, vol. **7**, p. 390; *W. L. Lewschin.* Journ. of Phys., 1939, vol. **1**, p. 265.
15. Ср. *А. Н. Севченко.* Тр. Гос. оптич. ин-та, 1941, т. **14**, стр. 65.
16. *S. I. Wawilow.* Zs. f. Phys., Bd. **53**, 1929, S. 665; *I. M. Frank u. S. I. Wawilow.* Zs. f. Phys., Bd. **69**, 1931, S. 100; *B. Sveschnikoff.* Acta physicochim. URSS, 1936, vol. **4**, p. 453; *S. I. Wawilow.* Acta physica polonica, 1936, vol. **5**, p. 417; *P. Debye.* Trans. Electrochem. Soc., vol. **82**, 1942, p. 265; *J. Q. Umberger* a. *V. K. La-Mer.* Journ. Am. Chem. Soc., vol. **67**, 1945, p. 1099; *E. W. Montroll.* Journ. of Chem. Phys., 1946, vol. **14**, p. 202.
17. *F. Perrin.* La fluorescence des solutions, Paris, 1925; *С. И. Вавилов.* ДАН СССР, т. **5**, 1936, стр. 271.
18. *I. M. Frank* u. *S. I. Wawilow.* Zs. f. Phys., Bd. **69**, 1931, S. 100.
19. *E. Gaviola* u. *P. Pringsheim.* Zs. f. Phys., Bd. **24**, 1924, S. 24; *F. Weigert* u. *G. Käppler.* Zs. f. Phys., Bd. **25**, 1924, S. 99; *W. L. Lewschin.* Zs. f. Phys., Bd. **26**, 1924, S. 274; *E. Gaviola.* Zs. f. Phys., Bd. **42**, 1927, S. 862; *W. Szymanowski.* Zs. f. Phys., Bd. **95**, 1935, S. 460.
20. *F. Perrin.* Ann. de Phys., 1932, vol. **17**, p. 283; *Ф. А. Душинский.* ДАН СССР, т. **14**, 1937, стр. 73.
21. *С. И. Вавилов.* ЖЭТФ, 1943, т. **13**, стр. 13.
22. *Ф. М. Пекерман.* ДАН СССР, т. **52**, 1946, стр. 409, 773.
23. *Ф. М. Пекерман.* ДАН СССР, 1947.
24. *С. И. Вавилов.* ДАН СССР, т. **42**, 1944, стр. 344.
25. *A. Jabłoński.* Zs. f. Phys., Bd. **95**, 1935, S. 53; *S. Szymanowski.* Zs. f. Phys., Bd. 95, 1935, S. 466.
26. *А. Н. Севченко.* ДАН СССР, т. **42**, 1944, стр. 349; Journ. of Phys., 1944, vol. **8**, p. 163.
27. *S. Venscateswaran* a. *A. Karl.* Indian Journ. of Phys., 1920, vol. **3**, p. 105; Zs. f. Phys. Chemie (B), 1928, Bd. **1**, S. 466.
28. *S. I. Wawilow* u. *L. A. Tumerman.* Zs. f. Phys., Bd. **54**, 1929, S. 270.
29. *П. А. Черенков.* ДАН СССР, т. 2, 1934, стр. 451.
30. *С. И. Вавилов.* ДАН СССР, т. 2, 1934, стр. 457.
31. *И. Е. Тамм* и *И. М. Франк.* ДАН СССР, т. **14**, 1937, стр. 107; *В. Л. Гинзбург.* ЖЭТФ, 1940, т. **10**, стр. 589.
32. *М. Д. Галанин.* ДАН СССР, 1947.

Академик

Н. Д. Папалекси

НЕЛИНЕЙНЫЕ КОЛЕБАНИЯ

За последние десятилетия развился и приобрел большое значение в науке и технике новый отдел учения о колебаниях — так называемые нелинейные колебания. Если вначале нелинейная трактовка колебательных проблем была достоянием лишь немногих ученых, работавших главным образом в области радиофизики, то за последние годы она получила широкое признание не только в научных, но и в инженерных кругах. Наряду с радио и акустикой, нелинейные колебания приобрели право гражданства также и в электротехнике, аэродинамике, теории авторегулирования и т. д. В развитии учения о нелинейных колебаниях большую роль сыграли работы наших ученых, в частности школы, связанной с именем академика Л. И. Мандельштама.

Известно, что когда две величины связаны между собой простой пропорциональностью, то зависимость между ними можно графически выразить прямой линией, почему она и получила название прямолинейной или просто л и н е й н о й зависимости. Линейную зависимость мы имеем, например, между удлинением и силой в законе Гука, между электрическим током и напряжением в простейшем законе Ома, между электродвижущей силой и изменением во времени силы тока в простейшем законе индукции и т. д. В более сложных случаях мы имеем линейную зависимость одной величины от ряда других величин, например законы Кирхгофа

26*

в разветвленных электрических цепях, содержащих постоянные сопротивления, самоиндукции и емкости. В еще более общем смысле мы говорим о линейных системах, если процессы в них передаются линейными дифференциальными уравнениями, обыкновенными и в частных производных, с коэффициентами, не зависящими от зависимых переменных. Так как математический аппарат линейных дифференциальных уравнений с постоянными коэффициентами сравнительно прост и в значительной мере, именно в связи с вопросами колебаний, он был математиками хорошо разработан, линейные колебания были с разных сторон и весьма детально исследованы в самых разнообразных областях знания и техники. При этом огромную роль в развитии теоретических методов расчета сыграла применимость принципа суперпозиции в области линейных колебаний, что привело к развитию весьма удобных символических (комплексных) методов анализа (Хевисайд, Карсон и др.).

Однако приведенные выше в качестве примера линейные зависимости далеко не всегда остаются линейными. Так, закон Гука, как известно, имеет ограниченную применимость—до так называемого предела пропорциональности. При дальнейшем увеличении действующей силы уже не будет пропорциональности между удлинением и силой. Точно так же закон Ома в простой форме не применим к проводникам, сопротивление которых зависит от силы тока (например, к тонким проволокам, нагреваемым током до высокой температуры, к ионизационным токам и токам через термоэлектронные приборы, к токам через индуктивности, содержащие железо, и т. п.). В этих случаях мы говорим о нелинейных зависимостях. Ввиду того, что процессы в системах, содержащих такие нелинейные элементы, как сопротивления, зависящие от силы тока, термоэлектронные приборы или самоиндукции с железом, а в механике — системы с сухим трением, не могут быть описаны линейными дифференциальными уравнениями, а во многих случаях приводятся к нелинейным дифференциальным уравнениям, обыкновенным и в частных производных, мы и называем такие системы нелинейными.

К системам нелинейным, как известно, не применим принцип суперпозиции, и их поведение часто весьма своеобразно и во многих отношениях совершенно непонятно с точки зрения линейных систем.

С другой стороны, математическая трактовка таких систем значительно менее развита, так как математическая теория нелинейных дифферен-

циальных уравнений встречает большие трудности и требует совершенно иного, более сложного, математического аппарата, чем линейные дифференциальные уравнения. Уже в случае статических проблем в нелинейных системах мы встречаемся с бо́льшими трудностями, чем в линейных системах. Например, в случае цепи, состоящей из постоянного сопротивления и плоского диода (кенотрона), к которому приложено постоянное напряжение, мы, как легко видеть, получаем для постоянного тока алгебраическое уравнение 3-й степени ввиду того, что зависимость между током i через диод и напряжением V имеет вид

$$i = kV^{3/2}$$

(формула Ленгмюра), т. е. вид полукубической параболы. В более сложных случаях нелинейности аналитическое решение приводится к трансцендентным уравнениям, так что обычно приходится прибегать к графическим методам решения.

Гораздо более сложно обстоит дело в случае динамических состояний. В простейшей статической задаче требуется определить зависимость между двумя величинами (например, между удлинением и приложенной силой в упругой системе или между силой постоянного тока и приложенной постоянной э. д. с.) как функцию параметров рассматриваемой нелинейной системы, характеризующих ее с физической стороны. В простейшем же случае воздействия на такую систему чисто синусоидального колебания результирующее динамическое колебание не будет, вообще говоря, синусоидальным, а более сложным, и для получения решения потребуется определить состав этого колебания, т. е. составляющие его частоты, их амплитуды и фазовые константы в зависимости от параметров, определяющих как самую нелинейную систему, так и воздействующее колебание. Следует подчеркнуть еще и следующее важное отличие нелинейных задач от линейных. Если в случае воздействия нескольких колебаний на линейную систему полное решение, в силу применимости принципа суперпозиции, получается как сумма отдельных решений для каждого из воздействующих колебаний отдельно, то в нелинейных системах принцип суперпозиции не применим, и в каждом конкретном случае приходится искать решение заново. При этом решения получаются более сложные, часто совершенно неожиданные, если подходить к ним с привычной для физиков и техников «линейной точки зрения».

Поясним это на классическом примере комбинационных тонов в акустике. Если предположить, как это обычно делается в механике, в частности в акустике, что в области малых колебаний существует пропорциональность между колебательным воздействием и вызываемым им смещением, то, в силу применимости принципа суперпозиции, мы должны были бы получить при воздействии на ухо двух гармонических колебаний лишь суперпозицию двух смещений тех же частот при отсутствии всяких других частот. Однако наблюдения показали, что при воздействии на ухо звуком, состоящим из двух произвольных простых тонов, ухо различает не только эти тона, но еще и кратные им, а также суммарный и разностный тона. Объяснение этому явлению было дано Гельмгольцем, заключившим, что причиной его является нелинейная зависимость между смещением s и воздействующей колебательной силой u. В самом деле, при простейшем предположении, что

$$s = \alpha u + \beta u^2 + \gamma u^3,$$

можно путем несложных вычислений легко убедиться в том, что воздействие сложного колебания

$$u = a_1 \sin \omega_1 t + a_2 \sin \omega_2 t$$

дает, наряду с частотами ω_1, ω_2, $2\omega_1$, $2\omega_2$, $3\omega_1$, $3\omega_2$, также и комбинационные частоты

$$\omega_1 + \omega_2,\ \omega_1 - \omega_2,\ 2\omega_1 - \omega_2,\ 2\omega_1 + \omega_2,\ \ 2\omega_2 - \omega_1,\ 2\omega_2 + \omega_1.$$

Уже из этого простого примера видно, какие сложные соотношения получаются в случае колебаний даже при простейших нелинейных зависимостях. По мере усложнения нелинейности системы трудности значительно возрастают как в математическом отношении, так и в смысле уяснения физической картины. Особенно интересные соотношения получаются для нелинейных колебательных систем, которые приобрели огромное значение для науки и техники. Одной из таких простейших систем, быть может старейшей, является математический маятник, колебания которого, как известно, управляются нелинейным дифференциальным уравнением

$$\frac{d^2\vartheta}{dt^2} + k \sin \vartheta = 0,$$

где ϑ — есть угол отклонения, а $k=\frac{g}{l}$ (g — ускорение силы тяжести, l — длина маятника). Для очень малых ϑ, когда $\sin\vartheta$ можно заменить через ϑ, уравнение движения переходит в линейное и легко решается. В общем же случае решение его представляет значительные трудности: оно приводит к эллиптическим функциям, и вместе с тем характер движения маятника сильно усложняется: колебания перестают быть гармоничными, период их зависит от амплитуды размаха (движение уже не изохронно). Такие колебания называются ангармоничными или псевдогармоничными, а сама система — псевдогармоничной. Много труда было затрачено математиками и физиками для создания теории маятника, что было очень важно для точного измерения времени в астрономии и геодезии. При этом были разработаны различные математические приемы, позволяющие с требуемой точностью получить решение задачи.

Другой, тоже старой и исключительно важной нелинейной колебательной проблемой, в которой силы также не пропорциональны расстоянию, является проблема периодического колебательного движения небесных тел солнечной системы. Как известно, строгого решения задачи для случая трех тел и более не существует и до сих пор, несмотря на то, что ею занимались со времен Ньютона крупнейшие математики мира. Однако для частного, практически весьма важного, случая, когда масса одного из трех тел значительно меньше массы других, как это имеет место для Луны в задаче Солнце — Земля — Луна, математикам удалось разработать ряд чрезвычайно ценных приближенных методов расчета, позволяющих с большой точностью вычислять движение Луны. К ним относится так называемый метод «возмущений», при котором влияние Луны рассматривается как возмущающий движение фактор; метод последовательных приближений, связанный с методом исключения «секулярных или вековых» членов; метод разложения в ряд по «малым параметрам», а также метод медленного изменения параметров. Развитием этих методов мы обязаны Эйлеру, Лапласу, Лагранжу, Пуассону, Тиссерану, Линдштедту, Пуанкаре и другим. Как мы увидим ниже, эти методы, разработанные в связи с проблемами небесной механики, оказались весьма пригодными при трактовке колебательных проблем теоретической физики (например, в теории атома и молекул), а особенно в специфических проблемах нелинейных колебаний, где они получили дальнейшее развитие.

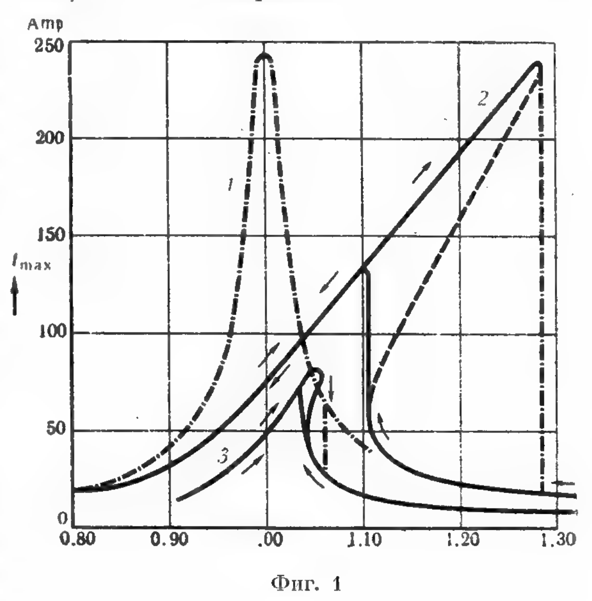
В рассмотренных нами в качестве примера нелинейных колебательных системах не было траты энергии, т. е., иначе говоря, эти системы были «консервативные». Однако действительные системы не консервативны и колебания в них постепенно затухают, если не будет компенсироваться расход энергии на поддержание в них длительных колебаний. Как известно, такие длительные или незатухающие колебания можно осуществить, воздействуя на систему извне колебательной силой, под действием которой система и совершает вынужденные колебания. Однако характер вынужденных колебаний в нелинейных системах, особенно колебательных, существенно отличается от характера вынужденных колебаний в линейных системах. Если, например, при воздействии гармонической силы на линейную систему мы получаем в ней гармоническое вынужденное колебание той же частоты, причем амплитуда его пропорциональна амплитуде внешней силы, то в нелинейных системах при таком воздействии вынужденное колебание будет, вообще говоря, состоять из целого спектра частот (кратных), а амплитуды отдельных гармоник будут сложным образом зависеть от амплитуды воздействующей силы. Еще более сложные зависимости получаются в нелинейных к о л е б а т е л ь н ы х системах при воздействии на них гармонической внешней силой: столь характерный и привычный для нас классический резонанс в линейных колебательных системах здесь сильно видоизменяется и во многих случаях совершенно утрачивает многие свои характерные черты.

Одним из наиболее характерных примеров, в котором особенно резко проявляются особенности резонанса в нелинейных колебательных системах, является так называемый «феррорезонанс», имеющий место в электрических колебательных цепях, содержащих индуктивности с железными сердечниками. На фиг. 1 приведены две феррорезонансные кривые (*2*, *3*) и кривая обыкновенного классического резонанса (*1*). В первом случае не только положение максимума вынужденных колебаний зависит сложным образом от амплитуды внешней э. д. с., но и форма кривой существенно различна в зависимости от того, как происходит настройка системы. Более того, само стационарное состояние определяется однозначно не параметрами системы, амплитудой и частотой вынуждающего колебания, а последовательностью предыдущих состояний, т. е. ее «предисторией». Такая «неоднозначность» стационарных состояний, помимо того, что она, разумеется, сильно осложняет расчеты, ставит вопрос о том, чем опреде-



Н. Папалекси

ляется физически наличие того или иного стационарного состояния, т. е. вопрос об их «устойчивости» или «неустойчивости». Необходимо заметить, что явления, подобные феррорезонансу, имеют место не только в практически важной области электротехники, но и в других областях физики и техники, которые, казалось бы, ничего общего с феррорезонансом не имеют. Так, А. А. Андронов и Г. С. Горелик недавно показали, что в циклотроне при движении релятивистских электронов, вследствие зависимости массы электрона от скорости, в стационарном режиме должны иметь место явления, подобные феррорезонансу.



Фиг. 1

Хотя с помощью внешнего периодического воздействия и можно получить желаемые длительные колебания (периодические и непериодические) как в линейных, так и в нелинейных системах, однако существуют и другие возможности генерации колебаний в автономных системах, т. е. в системах без внешнего воздействия. К ним относятся, например, часы с маятником и грузом, электромагнитные прерыватели (электрический звонок), питаемые электрическими элементами, духовые музыкальные инструменты (органная труба с мехом) и т. д. Возникшие в этих системах колебания благодаря наличию особых авторегулирующих устройств (обратной связи) могут длительно поддерживаться за счет местного источника энергии (потенциальной энергии груза, химической энергии электрических элементов, энергии воздуха, нагнетаемого в меха). Может быть, самой характерной особенностью этих систем (получивших название «автоколебательных» ввиду того, что в них колебания поддерживаются сами собой) является то обстоятельство, что в них устанавливается стационарное состояние в широких пределах независимо от величины

и рода усилия, которое выводит их из состояния покоя. Эта характерная особенность автоколебательных систем существенно отличает их не только от линейных консервативных колебательных систем, но и от нелинейных систем типа математического маятника, в которых стационарный режим зависит от начальных условий.

Исключительное значение получили автоколебания в связи с развитием радио, для которого проблема генерации непрерывных или незатухающих колебаний всевозможных частот, вплоть до самых высоких, приближающихся к инфракрасным, была одной из наиболее жизненных с самого его зарождения. Хотя уже генератор с электрической дугой Паульсена представлял собой мощный источник автоколебаний, однако только с появлением электронной радиолампы (Ли-де-Форест, 1907; Либен, 1911), ознаменовавшим фундаментальный переворот в развитии и принципа обратной связи (Мейснер, 1913), радио приобрело исключительное, можно сказать, почти идеальное средство не только для генерации колебаний (электромагнитных, механических, акустических и ультраакустических) всевозможных частот и формы, но и для усиления и вообще для управления формой и частотой колебаний.

Таким образом, помимо проблемы генерации колебаний с помощью автоколебательных систем, радио выдвинуло еще ряд других специфически нелинейных задач в области колебаний. Укажем, например, на проблему стабилизации частоты автоколебаний, а также на задачи трансформации частоты колебаний, их модуляции по амплитуде или частоте. Возникновение этих проблем не только во многих отношениях весьма значительно, а в некоторых даже фундаментально, расширило учение о колебаниях, но и потребовало разработки новых адэкватных методов теоретического анализа, а также более простых инженерных способов расчета колебательных процессов в новых нелинейных системах, создававшихся для решения практических задач.

Для того чтобы ясно было, с какими проблемами приходится иметь дело теории в этой новооткрывшейся области нелинейных колебаний, необходимо отчетливо себе представить, какие многообразные требования предъявляются радио к характеру и роду применяемых колебаний. Так, в некоторых случаях требуется генерация чисто синусоидальных колебаний неизменной частоты и амплитуды (например, для радиоинтерференционных методов навигации и точного измерения расстояний), в то время как

в других случаях необходимо иметь колебания очень сложного состава и формы (релаксационные колебания, прямоугольные и пилообразные развертки, применяемые в телевидении и радиолокации). Наряду с требованием сохранения неизменности формы колебаний иногда необходимо иметь быстрые изменения ее (модуляции при телефонии и телевидении). Такое многообразие, естественно, потребовало разработки новых методов математической трактовки возникающих здесь нелинейных задач, а часто и поисков новых путей подхода к ним.

Как уже было упомянуто выше, в связи с проблемами небесной механики были разработаны различные математические приемы нахождения приближенных решений нелинейных колебательных задач с любой степенью приближения. В основном эти методы касались нахождения стационарных (периодических или почти-периодических) колебательных режимов. Поэтому естественно было воспользоваться этими методами для теоретического рассмотрения стационарных режимов в нелинейных колебательных системах и в области радио. Естественно было также, что прежде всего подверглись анализу с помощью этих методов процессы в нелинейных системах, в которых колебания мало отличаются от синусоидальных, или, иначе говоря, в «слабо-нелинейных» системах.

Здесь необходимо отметить пионерскую роль известного физика и радиоспециалиста ван-дер-Поля. Им самим, а также совместно с известным радиофизиком Эпплтоном были впервые выяснены в ряде работ многие существенные стороны стационарных режимов в нелинейных колебательных системах с помощью впервые ими примененных к вопросам автоколебаний и другим нелинейным задачам радио различных методов небесной механики, а именно: метода последовательных приближений, теории возмущений, а также метода «медленно изменяющихся параметров», известного у нас под названием метода ван-дер-Поля. Так были получены приближенные выражения для стационарной амплитуды автоколебаний как при «мягком», так и при «жестком» режиме самовозбуждения, а также для «поправки на частоту»; дан качественный анализ вынужденных колебаний в автоколебательной системе (регенеративный приемник), в частности явления увлечения частоты (принудительной синхронизации). Необходимо заметить, что так как запросы практики неотложно требовали создания возможно более простых инженерных методов расчета автоколебательных радиотехнических

систем, то наряду с развитием физико-математического направления трактовки нелинейных колебательных задач стало развиваться и инженерно-техническое направление (Баркгаузен, Меллер, Иосс; у нас — А. И. Берг, Ю. Б. Кобзарев).

Хотя количественные методы, примененные ван-дер-Полем и другими, оказались весьма плодотворными на первом, пионерском, этапе развития исследований в области нелинейных колебаний, однако они смогли лишь частично удовлетворить запросам, предъявляемым теорией и практикой. Прежде всего их применение ограничивалось лишь теми случаями, когда колебательный процесс можно было в первом приближении аппроксимировать гармонической функцией, т. е. системами слабо-нелинейными, или так называемыми системами «томсоновского» типа. Для систем «не-томсоновского» типа (например, тормозные колодки, мультивибратор Абрагам — Блоха, релаксационные системы), а также сильно-нелинейных систем эти методы, вообще говоря, непригодны. Так, ван-дер-Полю, которому принадлежит заслуга первого систематического исследования релаксационных колебательных систем, приходилось для их трактовки в каждом отдельном случае прибегать к особым приемам, вплоть до графического способа решения дифференциальных уравнений. С другой стороны, применявшиеся первое время методы расчета не отличались достаточной строгостью и не были достаточно обоснованы, что не позволяло точно установить область их применимости, а также дать оценку точности получаемых результатов.

Большая заслуга в создании строгих адэкватных методов трактовки самых разнообразных нелинейных колебательных проблем принадлежит ученым СССР, и в первую очередь нелинейной школе физиков, связанной с именем академика Л. И. Мандельштама.

При рассмотрении любой колебательной проблемы, будь то автоколебания или вынужденные колебания, естественно возникают следующие, основные и для теории и для практики, вопросы. Прежде всего, конечно, необходимо определить характер стационарных колебаний, т. е. узнать, будет ли установившийся процесс периодический, или почти-периодический, какова будет форма стационарных колебаний, составляющие частоты, амплитуды и фазы. При этом, особенно в тех случаях, когда (как, например, при феррорезонансе) математически возможны несколько стационарных режимов, весьма существенно знать, какие из

них устойчивы и какова степень их устойчивости. Если возможны несколько устойчивых режимов, то часто бывает важно установить, в каких условиях получается тот или иной режим. Далее во многих задачах существенно знать, как протекает процесс установления колебаний и, в частности, как быстро они установятся. Мы не выделяем здесь особо существенного и важного вопроса о том, при каких начальных условиях могут возникнуть в данной системе колебания, так как этот вопрос, как более легкий, был раньше других достаточно строго и общо решен применением критериев Раутс — Гурвица, полученных еще раньше из рассмотрения устойчивости малых колебаний в механике.

Значительно труднее вопрос о нахождении стационарного режима. Для этой цели русскими физиками был детально разработан и широко применен так называемый метод малого параметра, развитый Пуанкаре для решения задач небесной механики. Основываясь на том, что решения дифференциальных уравнений изменяются непрерывно с изменением их параметров и начальных условий, Пуанкаре доказал, что если дифференциальное уравнение имеет для определенного значения данного параметра периодическое решение с периодом T, то для значений параметров, отличающихся на малую величину μ, решение также будет периодичным, причем в случае автономной системы (автоколебания) период будет мало отличаться от T, тогда как в случае неавтономной системы (вынужденные колебания) период будет точно равен T (или в особых случаях кратен T при так называемых периодических решениях второго рода, по терминологии Пуанкаре). Пуанкаре дал также метод получения решения в виде ряда по степеням μ.

С помощью метода малого параметра русским ученым не только удалось строго подтвердить результаты, полученные до того нестрогими методами, которые касались многих существенных сторон поведения нелинейных автономных (автоколебательных) и неавтономных систем, и полностью выяснить некоторые принципиально и практически важные спорные вопросы, но и открыть и изучить ряд новых замечательных явлений, значительно расширивших наши понятия о резонансе и оказавшихся весьма пригодными для использования в ряде практических случаев. Так, исходя из упомянутой выше возможности получения в нелинейных неавтономных системах периодических решений с периодом, кратным периоду действующей силы, т. е. периодических решений второго рода Пуанкаре,

академик Л. И. Мандельштам и автор обнаружили (1929), что в автоколебательной несамовозбужденной системе (например, в регенеративной ламповой системе, работающей в определенном режиме) можно возбудить вынужденные колебания с периодом, в два раза большим периода действующей внешней гармонической силы. Теория, развитая с помощью метода малого параметра, позволила охватить не только явления этого «резонанса второго рода», но и многие другие явления, как наблюдавшиеся раньше в несамовозбужденных автоколебательных системах под внешним гармоническим воздействием (синхронизация на обертоне — Кога), так и предсказанные этой теорией новые явления (явление тушения колебаний — Мандельштам и Папалекси), автопараметрические или дробные резонансы (изученные Ю. Б. Кобзаревым и В. В. Мигулиным), а также комбинационные резонансы (исследованные Б. М. Виленкиным и В. В. Мигулиным).

При изучении явлений резонанса второго рода Л. И. Мандельштамом и Н. Д. Папалекси было обращено внимание на то, что автоколебательные системы с внешним периодическим воздействием можно рассматривать как неавтономные колебательные системы с периодически изменяющимися параметрами. Ввиду того, что такое изменение параметров производится не непосредственно внешней силой, а через изменение независимых переменных, явление резонанса второго рода получило название «автопараметрического» резонанса в отличие от «параметрического» или «гетеропараметрического» резонанса, с которым оно имеет ряд общих черт и которое было изучено раньше.

С явлением параметрического резонанса физики и техники знакомы уже давно. Струна с периодически изменяющимся натяжением (опыт Мельде, 1861), качели, маятник с периодически изменяющейся длиной, вращающийся вал с асимметрично распределенными массами, электрическая колебательная цепь с периодически меняющейся емкостью или самоиндукцией (модулятор частоты в радио) — вот ряд примеров колебательных систем, в которых могут иметь место явления параметрического резонанса. Математически эти системы описываются дифференциальными уравнениями с периодическими коэффициентами. Как показывает теория (Рейли, 1883; Андронов и Леонтович, 1927), при определенном соотношении между частотой f изменения параметра и собственной частотой f_0 системы и при достаточной величине изменения параметра

возможно возникновение в системе колебаний на близкой к f_0 частоте f_1, удовлетворяющей соотношению

$$f/f_1 = 2/m,$$

где m — целое число. Особенно легко возникает этот эффект при соотношении частот $f = 2f_1$ (первый или главный параметрический резонанс). Математически явление параметрического резонанса приводится к наличию областей неустойчивых решений системы линейных дифференциальных уравнений с периодическими коэффициентами, которые — отчасти также в связи с проблемами небесной механики — были уже давно изучены математиками (Матье, Лиувилль и др.). Однако получение стационарных колебаний путем периодического изменения параметров системы — задача существенно нелинейная, так как только наличие нелинейности ограничивает беспредельное нарастание колебаний. Особый практический интерес получила теория параметрических колебаний (как принято кратко называть такие колебания) в связи с тем, что осуществленное Мандельштамом и Папалекси параметрическое возбуждение колебаний в электрической колебательной системе в отсутствие явных магнитных и электрических полей путем механического периодического изменения индуктивности (1930) и емкости (1933) системы привело к созданию генераторов переменного тока нового вида — так называемых параметрических генераторов (индуктивного и емкостного).

Применение метода малого параметра позволило дать приближенную теорию стационарного режима таких генераторов (Мандельштам и Папалекси, 1934). При этом рассмотрении, кроме нелинейных членов, малой величиной, порядка малости малого параметра, являлась относительная величина периодического изменения параметра (так называемая «глубина модуляции» параметра), а также и величина затухания системы, т. е. расход энергии в ней за период. Таким образом, здесь за исходную систему (в смысле Пуанкаре) бралась томсоновская колебательная система. Однако в практически важных случаях потребление энергии велико, и глубину модуляции нельзя считать малой, так что в качестве исходной системы нельзя брать консервативную томсоновскую систему. Поэтому оказалось необходимым распространить метод малого параметра на этот случай. Такая теория для конкретного случая системы с одной степенью свободы была дана в 1943 г. Мандельштамом и Папалекси, причем

в качестве исходного решения (нулевого) было взято (в духе Пуанкаре) периодическое решение линейного дифференциального уравнения с периодическими коэффициентами.

В связи с этими вопросами Мандельштамом была дана в 1942—1943 гг. и общая математическая теория метода малого параметра для решения системы дифференциальных уравнений с периодическими коэффициентами и «малой нелинейностью», что существенно расширило область применения этого метода.

Следует заметить, что метод малого параметра, плодотворность которого для получения стационарных решений чрезвычайно велика, получил дальнейшее развитие и в других направлениях. Так, Л. С. Понтрягин (1934) распространил его на автоколебательные системы, близкие к гамильтоновым, а Б. В. Булгаков (1940) — на системы, близкие к линейным, допускающие периодические решения, но не консервативные. Необходимо также отметить уточнение понятия «малого параметра» путем рассмотрения различных степеней μ, как положительных, так и отрицательных, которое позволило С. М. Рытову (1944) значительно расширить круг задач, решаемых методом малого параметра. С помощью такого уточнения ему вместе с сотрудниками удалось построить строгую теорию практически важной проблемы стабилизации частоты автоколебательных систем с помощью пьезоэлектрических и других резонаторов с малым затуханием.

В связи с обнаружением явления «асинхронного» возбуждения (Мандельштам, Папалекси, 1932; Кобзарев, 1933), заключающегося в возбуждении в несамовозбужденной системе колебаний, близких по частоте к самовозбужденным колебаниям системы независимо от частоты возбуждающей силы, метод малого параметра был распространен и на непериодические режимы. Здесь следует отметить исследования академика Н. М. Крылова и проф. Н. Н. Боголюбова, давших обоснование использования метода малого параметра для рассмотрения квази-периодических режимов.

Однако если метод малого параметра оказался весьма плодотворным для трактовки автономных и неавтономных систем с малой нелинейностью, как дискретных, так и с распределенными параметрами, то он все же мало пригоден для трактовки таких систем, в которых нелинейные члены не только не малы, но иногда больше других членов. Общих количествен-

ных методов пока еще не существует, однако во многих случаях здесь может быть с успехом применен так называемый «метод припасовки», примененный автором еще в 1911 г. для решения задачи выпрямления переменного тока в цепи с идеальным выпрямителем и самоиндукцией.

Сущность этого метода заключается в том, что входящая в задачу нелинейная зависимость аппроксимируется рядом прямолинейных отрезков (ломаной линией). Этим нелинейная задача разбивается на ряд линейных задач и решение сводится к отысканию постоянных интегрирования и времени из условия непрерывности процесса при переходе от одного прямолинейного участка к другому и из требования периодичности процесса. Такая идеализация приводит, однако, к представлению о «скачках» значений некоторых величин в стыках прямолинейных участков (например, направлении скорости, как это имеет место при упругом ударе шаров, или силы тока и напряжения на концах омического сопротивления), причем, как это уточнили Мандельштам и Папалекси, такие скачки могут испытывать лишь величины, изменение которых не связано с изменением энергии. Метод припасовки оказался очень полезным не только для исследования сильно-нелинейных автоколебательных систем (например, для теории электрического звонка, данной М. А. Леонтовичем в 1925 г.), но и для трактовки вынужденных колебаний в таких системах. Таковы, например, исследования Бен-Хартога о вынужденных колебаниях в системах с сухим трением, а также в релаксационных системах (Гуляев, Холоденко).

Особое значение приобрел метод припасовки в последнее время для решения задач автоматического регулирования. Как показали А. А. Андронов и его сотрудники, этот метод в комбинации с теорией точечных преобразований позволяет детально продискутировать не только периодические режимы, но и процессы установления во многих жизненных задачах автоматического регулирования, как то: в задаче Вышнеградского о теории прямого регулирования при наличии сухого трения (А. А. Андронов и А. Г. Майер, 1943), а также в некоторых задачах из области автоматического регулирования при наличии сервомоторов (А. Андронов и Н. Баутин и др.).

Очень ценный вклад в учение о нелинейных колебаниях был сделан нашими учеными и по вопросам устойчивости стационарных режимов. Еще в 90-х годах прошлого столетия (1892) нашим

знаменитым математиком А. М. Ляпуновым и одновременно французским математиком А. Пуанкаре были разработаны общие строгие методы решения вопросов динамической устойчивости периодических процессов, причем математически задача сводилась к определению «устойчивых» и «неустойчивых» областей решений системы линейных дифференциальных уравнений с периодическими или квази-периодическими коэффициентами, т. е. к решению рассмотренной выше задачи об условии возникновения параметрических колебаний. Нелинейной школе физиков, связанной с именем академика Л. И. Мандельштама, и в первую очередь А. А. Андронову и А. А. Витту (1930), принадлежит заслуга разработки на основе работ Ляпунова и Пуанкаре строгих методов трактовки динамической устойчивости нелинейных колебаний. С помощью этих методов, теперь общепризнанных, были полностью выяснены многие явления в нелинейных колебательных системах, связанные с многозначностью колебательных процессов в них.

Кроме условий возникновения колебаний, а также характера стационарных режимов и рода их устойчивости, большое теоретическое и практическое значение имеет знание переходных режимов, т. е. процессов установления и спадания колебаний. Как было указано выше, ван-дер-Поль пользовался для рассмотрения таких переходных процессов методом медленно изменяющихся параметров, физическая сущность которого заключается в том, что форма переходных колебаний рассматривается как синусоида с медленно изменяющимися во времени амплитудой и фазой. Этот метод, получивший у нас название метода ван-дер-Поля, оказался весьма адэкватным и чрезвычайно плодотворным и позволил не только изучить характер многих переходных режимов, но и получить приближенные выражения для стационарных режимов, как периодических, так и непериодических. Однако ван-дер-Поль не дал ни обоснования своего метода, ни, собственно говоря, общей рецептуры его применения. Им также не была определена область применимости этого метода. Л. И. Мандельштаму и Н. Д. Папалекси удалось в 1934 г. восполнить эти пробелы: дать математическое обоснование этого метода в применении к системам томсоновского типа, определить пределы его применимости в зависимости от величины малого параметра, а также указать правила получения так называемых «укороченных» уравнений. В дальнейшем метод медленно меняющихся параметров был распространен Г. С. Гореликом (1935) и

С. М. Рытовым (1936) и на нелинейные системы с периодически изменяющимися параметрами.

Чрезвычайно плодотворными и в известном смысле открывшими новую главу в исследовании нелинейных колебательных систем, в первую очередь автоколебательных, явились методы качественного анализа дифференциальных уравнений, описывающих эти системы. Открытие значения этих методов для теории нелинейных колебаний и превращение их в мощное орудие исследования — крупная заслуга А. А. Андронова и его многочисленных сотрудников и учеников. Отдельные вопросы, касающиеся связи между математической формой дифференциальных уравнений, описывающих данную систему, и физическими свойствами колебаний в ней, рассматривались и раньше (Картан, 1925; Льенар, 1928), однако лишь А. А. Андроновым было обнаружено, что можно на основе качественной теории дифференциальных уравнений, развитой Пуанкаре, исключительно наглядным образом изобразить на так называемой фазовой плоскости поведение колебательной системы и дать точное математическое определение автоколебаний. Пользуясь понятиями особых точек, предельных циклов и сепаратрисс, можно, как показал Андронов, из топографической картины на фазовой плоскости (фазовых траекторий) получить ответ на вопросы об устойчивости начального состояния системы, о характере возможных стационарных режимов, а также об их устойчивости. Исключительная наглядность метода «фазовой плоскости» и то обстоятельство, что область его применимости не ограничена ни требованием малой нелинейности, ни типом колебаний (он применим как к системам томсоновского типа, так и к релаксационным), делают его чрезвычайно ценным орудием исследования в самых разнообразных областях нелинейных колебаний. Вводя понятие «скачка» непрерывности, можно распространить этот метод качественного анализа и на «разрывные» колебания, например на колебания в мультивибраторе Абрагам — Блоха, исследованные Андроновым и Виттом в 1929 г. В случае малого параметра можно, пользуясь тем, что укороченные уравнения, к которым приводит метод ван-дер-Поля, не содержат явно времени, применить метод фазовой плоскости и к неавтономным системам (А. А. Андронов и А. А. Витт, 1930 г.).

Теория качественного анализа дифференциальных уравнений путем исследования фазовых траекторий в применении к колебательным проб-

27*

лемам была до сих пор детально разработана лишь для систем с одной степенью свободы, т. е. для фазовой плоскости, где она дает особенно наглядные результаты. Однако принципиально она может быть применима к системам с любым числом степеней свободы. В последнее время Андронову и Баутину (1944) удалось использовать метод качественного анализа в трехмерном пространстве, где он был ими с успехом применен для решения ряда задач автоматического регулирования.

Наряду с рассмотренными выше общими методами анализа нелинейных колебательных проблем (метод малого параметра, метод медленно меняющихся параметров или метод медленных возмущений, метод припасовки и метод фазовой плоскости — фазового пространства), являющимися основными для трактовки нелинейных колебательных систем, как дискретных, так и с распределенными постоянными, в ряде случаев целесообразно применять и другие математические приемы. Так, для решения колебательных задач в системах особого класса, а именно таких, взаимодействие между которыми определяется не мгновенным их состоянием, а зависит от их предшествующих состояний, так сказать, от их предистории, целесообразно, как показал В. М. Бовшеверов (1936), пользоваться методом «итераций», который был развит Кенигсом и Лемерэ. К этому классу относятся так называемые системы с «запаздывающей» обратной связью, например схема Куликова — Шиловского (1928) для измерения расстояний с помощью радиоволн, эхолот Жакэ — Бодена, органные трубы, некоторые типы сервомеханизмов и некоторые схемы генерации дециметровых волн. Метод итераций был также с успехом применен А. А. Виттом для исследования колебаний скрипичной струны, приводимой в звучание смычком, лехеровой системы, возбуждаемой при помощи электронной лампы, а также для исследования поведения релаксационных автоколебательных систем под внешним воздействием (Гуляев, Холоденко, Теодорчик и др.).

Рассмотренные методы математической трактовки были во многих случаях с успехом применены и к системам с распределенными постоянными. Выше уже упоминалась исследованная Виттом лехерова система, возбуждаемая при помощи электронной лампы. В последнее время С. М. Рытов и М. Е. Жаботинский (1946) применили с более общей точки зрения метод малого параметра к нелинейным системам штурм-лиувиллевского типа. С помощью этих методов были также проанализированы

некоторые вопросы колебаний в новых электронных приборах для генерации очень коротких электромагнитных волн (магнетронах и клистронах). Сюда относятся также и некоторые нелинейные эффекты при распространении радиоволн в атмосфере, например люксембург-горьковский эффект, исследованный как зарубежными, так и нашими учеными (Мартин, Бейль, ван-дер-Поль, С. М. Рытов и др.). Необходимо, однако, подчеркнуть, что хотя в изучении колебательных процессов в нелинейных системах с распределенными постоянными и были достигнуты значительные результаты, однако эта область нелинейных колебаний остается еще во многих отношениях неисследованной, главным образом из-за того, что она математически еще очень мало разработана. Здесь следует отметить существенный вклад, внесенный в эту область приближенным методом академика С. А. Чаплыгина.

С помощью очерченных выше основных методов теоретического анализа в последние десятилетия были детально изучены разнообразные нелинейные колебательные системы в различных областях физики и техники, причем часто нелинейные эффекты, обнаруженные и исследованные в одной из областей физики, например в области электрических колебаний, были затем выявлены и детально изучены в области механических и акустических колебаний. Так, явление «принудительной» синхронизации, известное в механике еще со времен Гюйгенса, наблюдавшего взаимную синхронизацию часов, подвешенных на одной и той же стене, и широко используемое в радио, было затем обнаружено и исследовано С. Э. Хайкиным и К. Ф. Теодорчиком и в акустике. С. П. Стрелков исследовал автоколебания маятника Фроуда, представляющего собой простейший механический аналог динатронного лампового генератора. Хайкин и Кайдановский осуществили и исследовали механические автоколебательные явления, обусловленные сухим трением и проявляющиеся, например, в виде д р е б е з ж а н и я т о р м о з о в. На электрической модели сердца ван-дер-Поль показал, что к релаксационным колебаниям относится также деятельность нашей сердечной мышцы.

Исследования последних лет выявили также огромную роль, которую играют нелинейные колебания в автоматике, авиации и электротехнике. Значительная часть нелинейных колебательных проблем, встречающихся в различных областях техники, особенно в автоматике, связана не с генерацией автоколебаний, как в радио, а с устранением вредных

автоколебаний, возникающих часто при наладке установок с авторегулирующимися механизмами, а также разного рода следящих систем, например автопилотов и авторулей. Многие из относящихся сюда вопросов авторегулирования рассмотрены в работах как наших ученых (Андронов, Баутин, Булгаков, Горелик и др.), так и иностранных (Синж — Канада, Минорский — США). Сюда относятся также и вопросы автоколебаний, поддерживаемых аэродинамическими силами (Стрелков; Стрелков и Аронович). В электротехнике к этому классу проблем относятся вопросы устойчивости работы как отдельных электрических машин, например «качание» мотора постоянного тока (Жане, 1925; З. и Г. Картан, 1928), автоколебания синхронного мотора (Дрейфус, 1911; Роговский, 1915; Оллендорф, 1933; Власов, 1939, и др.), так и параллельной работы двух синхронных машин (Н. М. Крылов и Боголюбов, Горев и др.).

Как видно из этого краткого и поэтому далеко неполного очерка, роль нелинейных колебаний в самых разнообразных областях знания и техники в последнее время чрезвычайно возросла. Учение о нелинейных колебаниях из собрания отдельных, часто любопытных, а иногда и парадоксальных частных примеров превратилось в большой самостоятельный раздел науки, влияние которого на развитие ряда разделов техники, уже и теперь весьма заметное, все возрастает. В развитии области нелинейных колебаний видная роль принадлежит ученым нашей страны как в заложении фундаментальных математических основ (Ляпунов, Жуковский, Чаплыгин), так и в открытии новых явлений и в создании и разработке новых адэкватных методов их трактовки. Если в первый период развития учения о нелинейных колебаниях, вызванного появлением электронной лампы, ведущая роль в исследовании нелинейных колебательных систем и развитии их теории принадлежала ученым других стран, то, начиная с 1927/28 г., она постепенно перешла к нам, и в настоящее время ведущая роль наших ученых в области нелинейных колебаний получила всеобщее признание.

Академик
В. В. Шулейкин

ТЕПЛОВЫЕ ВЗАИМОДЕЙСТВИЯ МЕЖДУ ОКЕАНОМ, АТМОСФЕРОЙ И МАТЕРИКОМ

Еще недавно, каких-нибудь десять лет назад, оглядываясь на путь, пройденный советской физикой моря, работники этой области науки могли с удовлетворением отмечать, что различные проблемы динамики течений и волн, казавшиеся прежде разрозненными, нацело спаялись между собой идейной связью и одновременно оказались связанными с глубокими и интересными проблемами термики моря. Сама термика моря в нашем понимании утратила тогда облик какой-то главы статики океана, как значилась она в старых фундаментальных книгах, выходивших на различных языках.

Сейчас, к тридцатой годовщине великого исторического рубежа, итоги нашей науки позволяют нам с еще бо́льшим удовлетворением установить, что почин советских мореведов в этой области точного естествознания не только привел к широкому развитию во всех странах учения о тепловом балансе моря, но и позволил связать количественными зависимостями элементы теплового баланса моря с поведением атмосферы над морем и над материками.

Полтора десятка лет назад, исследуя тепло, поступающее на поверхность моря от Солнца, обратное излучение этого тепла, потерю его на испарение морской воды и на подогревание холодного воздуха, а также исследуя тепло, выделяющееся при замерзании воды и отнимаемое от

окружающих вод при таянии льда, мы обнаружили, в частности, что все наши окраинные северные моря — и Баренцово, и Карское, и Восточно-Сибирское, и море Лаптевых — расходуют громадное количество тепла, когда их поверхность оказывается на короткий срок свободной от ледяного покрова. С точки зрения нашей теории, такое временное освобождение от льда оказывается обязанным лишь поступлению чрезвычайно большого количества тепла от глубинных теплых течений, входящих в эти моря с севера, из бассейна собственно Ледовитого океана.

За истекшие годы это предположение подтверждалось на каждом шагу; вдобавок к теплым струям, входящим в Баренцово море и в северо-западную часть Карского, неподалеку от Земли Франца Иосифа, в 1935 г. восточной экспедицией на ледокольном пароходе «Садко» были открыты мощные теплые струи. Замечательные результаты по исследованию глубинных теплых вод в Ледовитом океане были получены дрейфующей станцией «Северный Полюс». Наконец, практики-прогнозисты, исходя из идеи о решающем значении осеннего расхода тепла (до возникновения толстого ледяного покрова), научились предсказывать навигационные условия на некоторых участках Северного морского пути по осеннему и раннему зимнему режиму предшествовавшего года.

Все это укрепило уверенность в справедливости теории теплового баланса и дало толчок к распространению ее на сложные тепловые взаимодействия между океаном и атмосферой, с одной стороны, и между атмосферой и материками — с другой.

Очень полезным оказался метод выделения влияния океана на климат материка, основанный на построении температурных изаномал: этим устранилось маскирующее действие потоков тепла, которые идут вдоль меридианов от более теплых широтных зон к более холодным. На фиг. 1 изображена карта температурных изаномал января для Европы. Кривые соединяют между собой точки, в которых средняя температура воздуха в январе отличается на то или иное число градусов от осредненной по всей параллели, соответствующей широте данной точки. Чем ближе к Атлантическому океану лежат точки материка, тем выше температурная «аномалия» в этих точках, тем больше превышает температура воздуха в них осредненную широтную температуру. В свою очередь, осредненные широтные температуры позволили бы судить о кли-

матических различиях, не связанных с океаном и обусловленных лишь различием в солярном режиме и переносом тепла в меридиональном направлении.

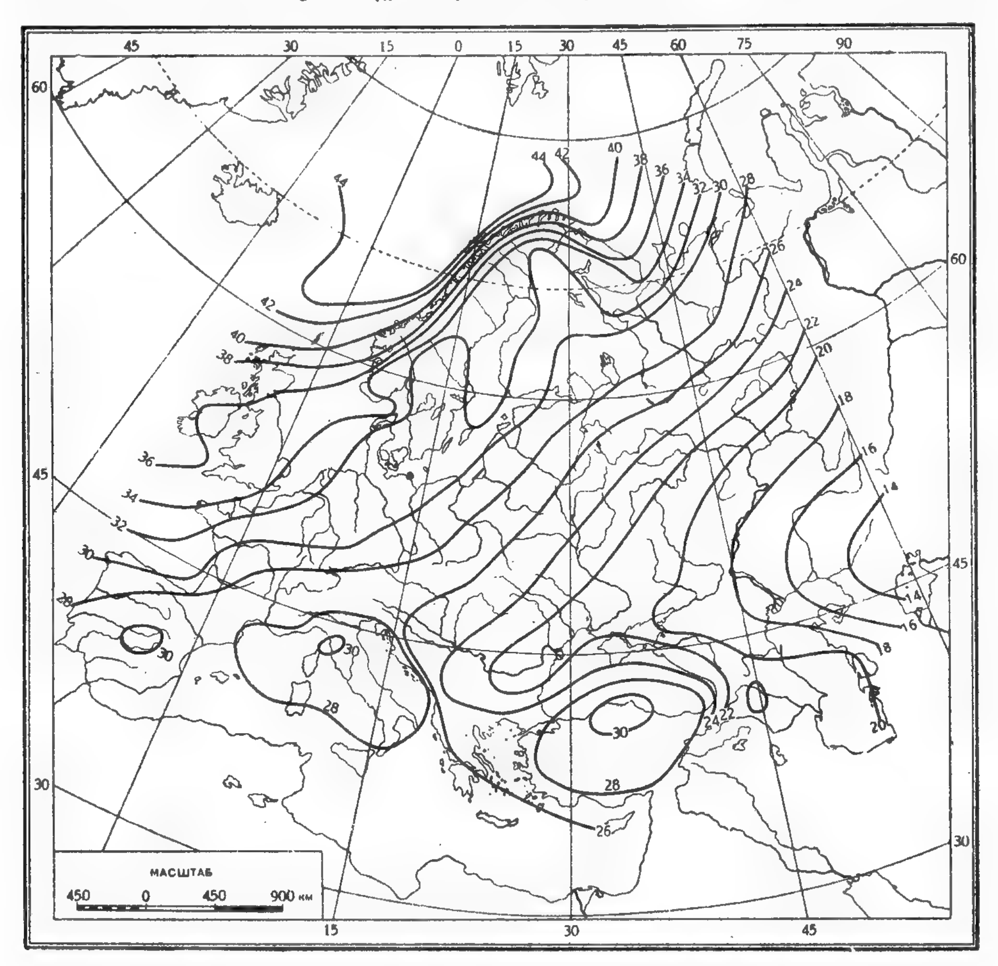
Вводя понятие о некотором формальном коэффициенте теплопроводности атмосферы на путях с океана на материки, оказалось возможным подсчитать количество тепла, которое переносится в зимнее время с океана на материк над единицей длины береговой черты. Также оказалось возможным вычислить количество тепла, которое заимствуется от подобных тепловых потоков, выделяясь в нижних слоях атмосферы как некоторое дополнительное по отношению к теплу, поступающему непосредственно от солнечных лучей.

Результаты таких вычислений, проделанных, в частности, по отношению к Ленинградской области, полностью совпали с результатами работ Павловской аэрологической обсерватории по тепловому балансу атмосферы в той же области: то количество тепла, которое, согласно нашей теории, выделяется потоками морского происхождения и идет на дополнительное подогревание атмосферы, в точности покрыло «тепловой дефицит», обнаруженный работниками аэрологической обсерватории близ Ленинграда.

Еще более убедительное подтверждение теории обнаружилось в области динамики атмосферы, подогреваемой над морем и охлаждающейся над материком. Именно, расположение температурных изаномал (фиг. 1) неминуемо должно быть связано с совершенно аналогичным расположением климатологических изобар в зимнем муссонном поле: повышенному прогреву деятельного слоя тропосферы должно соответствовать пониженное атмосферное давление и, наоборот, пониженному подогреву воздуха должно соответствовать повышенное давление. Количественно это выражается уравнением, которое следует рассматривать как основное уравнение муссонного поля: оно связывает изменение атмосферного давления на протяжении единицы длины в горизонтальном направлении с соответствующим изменением температуры воздуха. Его можно записать в следующем простом виде: градиент давления $= \Pi \times$ градиент температуры, причем Π, выраженное в абсолютных единицах, численно равно 1.6×10^6.

В таком случае, если наша теория верна, то карта температурных изаномал Европы (фиг. 1) должна быть весьма схожа с картой климато-

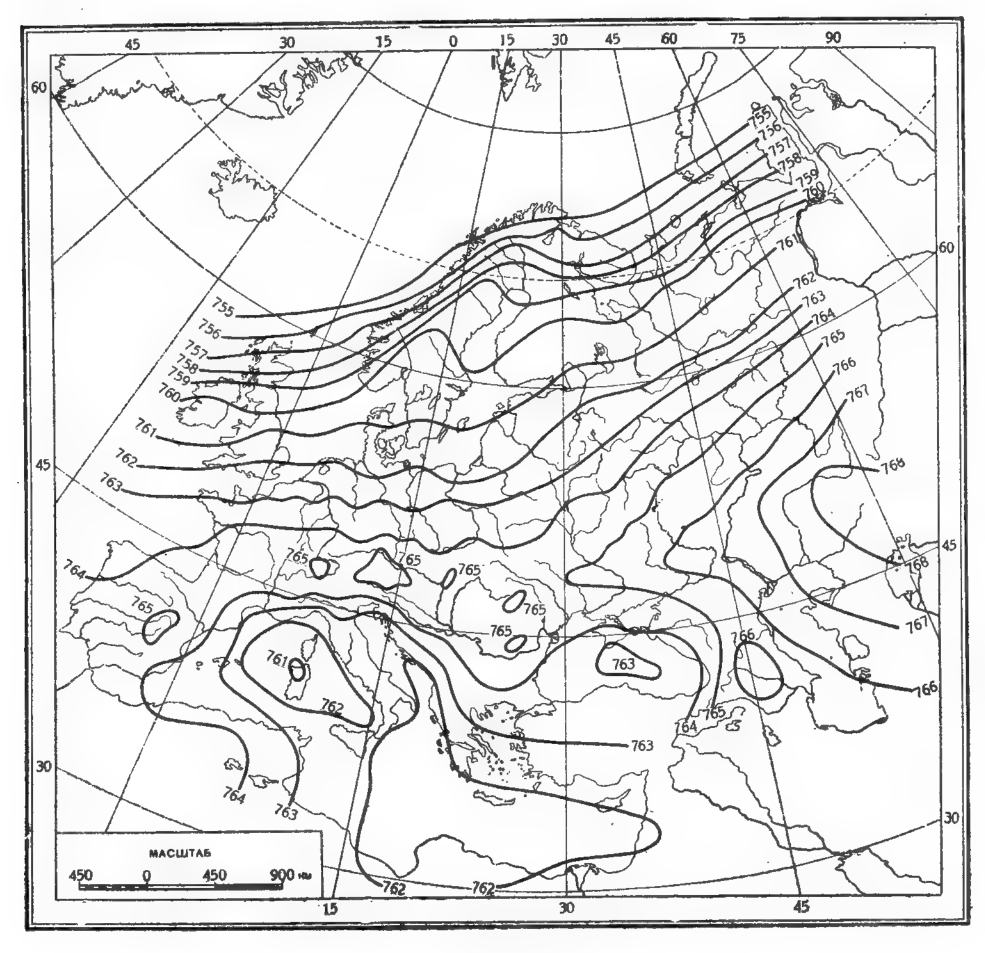
логических изобар Европы, построенных также для января (фиг. 2). Сопоставляя ее с картой (фиг. 1), можно убедиться в большом сходстве



Фиг. 1

между ними, вплоть до отдельных деталей. В самом деле, на обеих картах отмечаются совершенно аналогичные «языки», вторгающиеся из Азии в Европу между Балтийским и Белым морями, с одной стороны, и Черным и Каспийским морями — с другой. Как ни скромны размеры двух

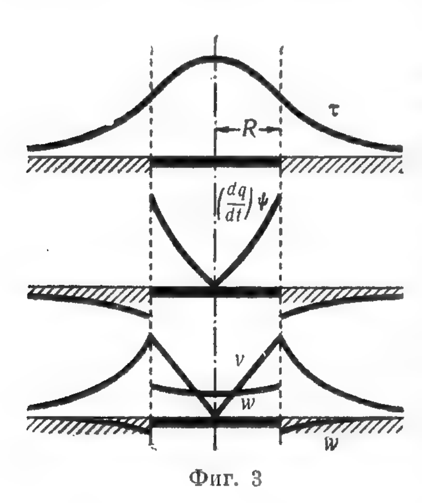
последних морей, обе карты показывают явное обострение признаков «материкового климата» при удалении от них на север или северо-запад.



Фиг. 2

При переходе через некий «теплораздел», совпадающий с линией Ворошиловград — Павлоград — Киев, дальнейшее продвижение на север и северо-запад сопровождается новым повышением влияния моря — на этот раз Баренцова, Гренландского, Балтийского.

Особо надо подчеркнуть, что барометрическое давление является показателем состояния всей толщи атмосферы, тогда как температуры, измеряемые на наземных или на корабельных станциях, свидетельствуют лишь о режиме нижних слоев воздуха. Замечательное сходство двух карт (фиг. 1 и 2) указывает на возможность суждения о важнейших процессах в атмосфере по поведению температуры нижних слоев атмосферы. Столь же справедливо и другое заключение: исследование схематизированной условной «теплопроводности» всего деятельного слоя тропосферы может дать представление о дополнительном прогреве воздуха над морем и о постепенном уменьшении подобного дополнительного прогрева воздуха при удалении соответствующих географических точек от моря. Разумеется, одновременно с такой условной теплопроводностью атмосферы необходимо рассматривать еще два взаимно противоположных процесса: поступление тепла в атмосферу от поверхностной морской воды, соприкасающейся с воздухом и обладающей (в зимнее время) более высокой температурой, и излучение тепла в межпланетное пространство.



Фиг. 3

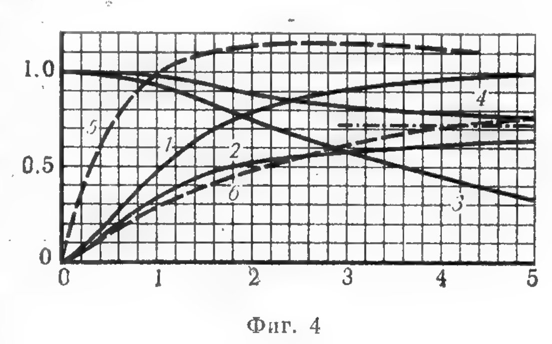
Задача решается очень просто для двух идеализированных случаев: для круглого моря, возникшего среди безгранично простирающегося материка, и для круглого материка, обрамленного безгранично простирающимся океаном.

На фиг. 3 изображена схема муссонного поля для первого варианта. Черная полоса схематически изображает разрез моря вертикальной плоскостью. По обе стороны от него обозначены штриховкой примыкающие к нему части материка. В вертикальной плоскости сечения лежит ось симметрии, нанесенная чередующимися черточками и точками и проходящая через центр круглого моря. На этой плоскости, прежде всего, видна холмообразная кривая τ, выражающая закон распределения температуры над морем и над материком. Температура здесь отсчитывается от некоторого условного нуля: от той температуры, которую принял бы

воздух, если бы земной шар полностью обсох и в зимнее время полностью отсутствовало бы дополнительное подогревание воздуха поверхностной морской водой.

Кривая $\left(\frac{dq}{dt}\right)\psi$ обозначает изменения количеств тепла, получаемых воздухом с единицы поверхности моря в морском районе или, соответственно, затрачиваемых атмосферой на дополнительное излучение над материком, где это дополнительное тепло черпается от тепловых потоков морского происхождения в воздухе.

Наконец, кривая v схематически изображает распределение ветров муссонного происхождения; она ясно показывает, что эти ветры должны достигать наибольшей силы в прибрежных районах.



Фиг. 4

На другой диаграмме (фиг. 4) изображена зависимость теплового и ветрового режима над морем, материком или над береговой чертой от р а з м е р о в самого круглого моря, обрамленного материком, или же круглого материка, обрамленного океаном.

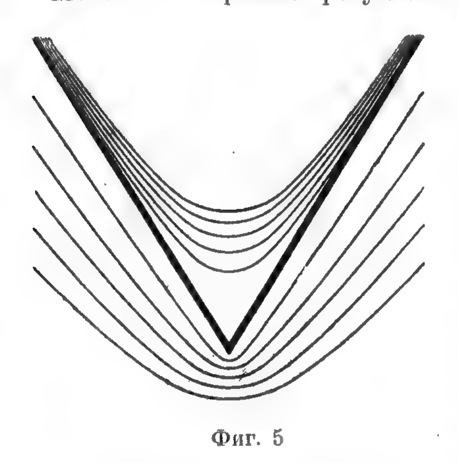
Кривая *1* показывает, как увеличивается прирост температуры (над упомянутым условным нулем) в центральной части моря при увеличении радиуса моря. Значения этого радиуса (в тысячах километров) отмечены под горизонтальной осью диаграммы. Кривая *2* выражает изменение температурного режима над берегом; кривая *3* рисует падение температуры над центральной областью материка в зависимости от размеров последнего. Исходной температурой, принимаемой за единицу, считается та, какую приобрел бы воздух, если бы в е с ь земной шар оделся океаном. Кривая *4* выражает падение температуры над берегом материка, размеры которого нарастают.

Разумеется, при безграничном нарастании размеров, и кривая *2*, и кривая *4* стремятся к некоторому общему пределу; этот предел соответствует полупространству, занятому морем, протянувшимся в бесконечность в одну сторону от береговой линии, и полупространству, занятому

материком, протянувшимся аналогично в другую сторону от берега.

Любопытно, что все эти кривые, вычисленные по нашей теории, хорошо совпадают с данными непосредственных систематических измерений температур над океанами, внутренними морями, материками, островами и их берегами. Вычисленные температуры очень близки к тем, которые получаются из наблюдений при соответствующих размерах моря или материка.

Столь же хорошие результаты дают вычисления скоростей ветра, ожидаемых на берегах моря, обрамленного материком (кривая *5*), или материка, обрамленного океаном (кривая *6*).



Фиг. 5

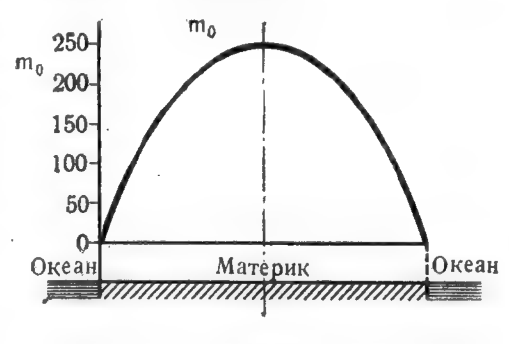
Но, вообще говоря, форма морей и форма материков отличается от той идеализированной круглой, лежавшей в основе всех вычислений. Следовательно, большой интерес представляет выяснение роли ф о р м ы самой береговой линии, отделяющей воду от суши. В этом направлении теория также привела к очень существенным результатам. Выяснилось, что температурное и барическое поле в море против остроконечного мыса или на материке против остроконечного залива неизменно должно о б о с т р и т ь с я в согласии со схемой, изображенной на фиг. 5. Напротив, как видно на той же схематической диаграмме, вычисленной теоретически, на протяжении самого мыса или на протяжении самого залива муссонное поле оказывается менее напряженным. Обо всем этом свидетельствуют изолинии как против мыса (залива), так и на его протяжении (фиг. 5).

С этой точки зрения совершенно естественными являются постоянные ураганные штормы, которыми славятся и мыс Горн и мыс Доброй Надежды (прежде носивший более подходящее название — мыс Бурь). Даже скромные мысы Крымского полуострова, вклинивающиеся в Черное море (Тарханкут, Сарыч, Херсонесский и др.), не говоря уже о Канином Носе и о м. Желания на северной оконечности Новой Земли, всегда

порождают повышенные скорости ветра над морем по соседству с ними. Любопытно, что подобное же усиление ветров, в полном согласии с теорией, наблюдается против острых оконечностей озера Байкал, врезающихся в громадный материк Азии.

Но выводы теории муссонного происхождения климата простираются еще дальше и позволяют проследить за очень интересными особенностями движения всей Земли как некоторой самостоятельной механической системы.

В самом деле, значительное повышение атмосферного давления над материками и понижение его над океаном в зимнее время свидетельствуют о том, что с океана на материки перетекает к зиме громадное дополнительное количество воздуха, находившегося в летнее время над океанами. Исследуя механизм воздухообмена между океаном и материками в муссонных потоках, легко притти к выводу, что дополнительные массы воздуха располагаются над неким идеализированным круглым материком так, как это изображено на диаграмме фиг. 6. Над каждой единицей поверхности материка зимой оказывается дополнительная масса воздуха, равная, примерно, около одной четверти тонны. Подсчет показывает, далее, что над всем соединенным Европейско-Азиатским материком по той же причине появляется дополнительная масса воздуха, достигающая 300 000 000 000 000 тонн. Значительно меньшие, но все же довольно существенные дополнительные массы воздуха оказываются расположенными зимой над материком Америки. Совершенно очевидно, что соответствующее количество воздуха оказывается зимой изъятым из слоя атмосферы, залегающего над океанами.



Фиг. 6

Но ведь океаны и материки расположены на земной поверхности несимметрично по отношению к оси вращения Земли. Следовательно, попеременный переход воздушных масс — то с океанов на материки, то с материков на просторы мирового океана — должен

приводить к непрерывным изменениям моментов инерции соответствующих масс относительно оси вращения.

Из законов механики следует, однако, что подобная система не может вращаться совершенно спокойно: если бы Земля была каким-то маховым колесом, ось которого лежала бы в неподвижных подшипниках, то такой маховик был бы обязан бить — нецентрированные, несимметрично расположенные массы привели бы к возникновению больших давлений на ось, перпендикулярно к последней.

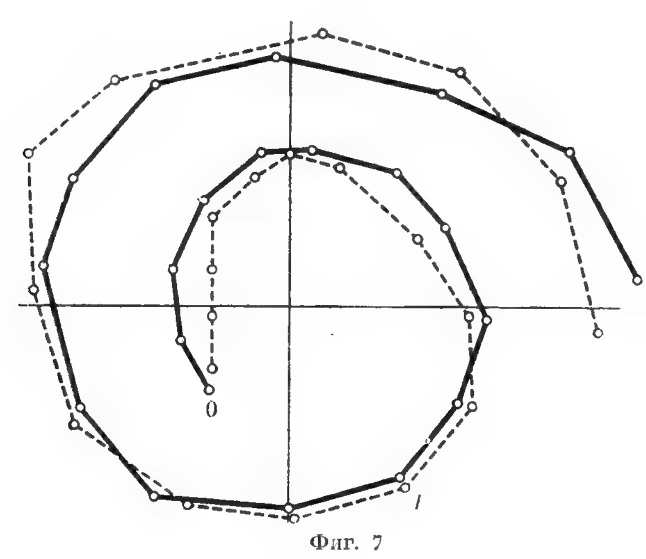
Земной шар, в отличие от такого маховика, предоставлен самому себе: он вращается без всяких неподвижных подшипников. По этой причине законы механики требуют непрерывного смещения самой оси вращения внутри тела Земли в зависимости от смещения воздушных масс с океана на материки и обратно.

Наши исследователи определили, как именно должна смещаться земная ось вращения по заданным смещениям дополнительных воздушных масс. Тем самым они определили, как должны блуждать около своих средних положений оба полюса вращения, и северный и южный.

В нашей же стране был суммирован опыт астрономов всех стран, наблюдавших в действительности блуждания земных полюсов вращения по изменениям высот светил, вычисленных для соответствующих широт и соответствующих моментов времени. Оказалось, что вычисленные перемещения полюсов очень хорошо совпадают с наблюденными, в чем не трудно убедиться, взглянув на фиг. 7. На ней пунктиром изображен наблюдаемый путь, по которому смещался северный полюс Земли за некоторый промежуток времени, а сплошной линией — путь полюса, вычисленный по нашей теории.

В той же работе Черноморской гидрофизической станции Академии Наук СССР были попутно вскрыты также и другие важные обстоятельства, позволившие объяснить до той поры необъясненные явления. Прежде всего, до последнего времени было непонятно, почему в некоторые годы земные полюсы смещаются по отношению к своему среднему положению очень слабо, в другие же годы смещения их достигают значительно бо́льших размеров. Оказалось, что это объясняется непостоянством протяжения «геофизического года», т. е. промежутка между соседними зимами: в один год зима может оказаться ранней, а в следующий — поздней, а потому геофизический год будет заметно превышать 12 меся-

цев. Таким образом, геофизический год может оказаться весьма близким к тому промежутку времени, который носит название «периода свободной нутации Земли». Можно показать, что при этом неминуемо должна повыситься амплитуда отклонений полюсов от их средних положений, подобно тому как возрастает амплитуда при всяком ином явлении резонанса.



Фиг. 7

Анализируя карту распределения дополнительных воздушных масс, построенную по новейшим данным, полученным в цитированной работе, оказалось возможным обнаружить границы между областями Азии и Европы, находящимися под преобладающим влиянием зимнего муссона, и областями, в которых преобладает влияние летнего муссона. Между этими двумя областями залегла ярко выраженная пограничная полоса, протянувшаяся, примерно, от Берингова пролива через оз. Байкал, Каспийское, Черное, Средиземное моря до Пиренейского полуострова. Области, находящиеся к юго-востоку от этого рубежа, испытывают на себе преимущественное влияние летнего муссона; зимний муссон в них выражен слабо (как известно, в Индийском океане, у берегов Азии, зимой часто стоит штилевая погода). Напротив, области, расположенные к северо-западу от того же рубэжа, обязаны своим климатом преимуще-

ственно зимнему муссону; летний там выражен очень слабо (в Баренцовом море, например, летом преобладают совсем слабые и умеренные ветры).

Отмеченный «муссонораздел» на соединенном Евразийском материке был обнаружен благодаря тому, что на карту наносились не изаллобары, приведенные к уровню моря (как обычно), а изаллобары, приведенные к высотам соответствующих точек земной поверхности. Этим обеспечивалось определение разностей атмосферных давлений между январем и июлем, исправленных на толщину слоя, занятого в пространстве не воздухом, а горными массивами. По исправленным разностям давлений контролировались теоретически вычисленные дополнительные массы воздуха.

Обнадеживающие результаты, полученные при построении основ физической теории климата, внушили намерение предпринять попытки к построению хотя бы приближенной и грубой физической теории погоды. За основу были приняты следующие положения:

а) Общая циркуляция атмосферы и общая циркуляция океанических вод обусловлены работой каких-то «тепловых машин», черпающих энергию от Солнца и затем излучающих тепло в мировое пространство.

б) Подобно всякой иной тепловой машине, превращение тепловой энергии в механическую происходит здесь при обязательном наличии нагревателей и холодильников.

в) Машины одного рода, осуществляющиеся в атмосфере, работают благодаря переходу тепла от нагревателя — тропических зон к холодильникам — полярным зонам земного шара.

г) Машины другого рода, иногда еще резче выраженные в атмосфере, работают между нагревателем и холодильником, которые в зависимости от времени года могут меняться местами: зимой роль нагревателя исполняет мировой океан, а роль холодильника — материки; напротив, летом материки служат нагревателями, а океан играет роль холодильника.

д) Машины первого рода приводят к так называемой зональной и отчасти междузональной циркуляции атмосферы; работа их сопряжена с переносом тепла вдоль меридианов.

е) Машины второго рода проявляются в муссонной циркуляции между материками и океаном, а также между материками и отдельными (иногда—

внутренними) морями; их работа сопряжена с переносом тепла с моря на материк зимой или с материка на море летом.

ж) Некоторый средний режим этих тепловых машин мы воспринимаем как климат. К сожалению, научившись вычислять тепловые характеристики климата, мы еще не умеем производить подобных вычислений по отношению к влажности воздуха, которая, несомненно, связана с режимом наших тепловых машин в атмосфере.

з) Но вообще режим их работы непостоянен, как непостоянен, строго говоря, режим всякой машины, не снабженной сколько-нибудь точным регулятором. Колебания этого режима мы и воспринимаем как смену погоды.

Не трудно разобраться в характере тех колебаний, которые могут возникнуть в системе тепловых потоков, охватывающих тропосферу и проносящихся как в меридиональном направлении, так и в направлениях, связывающих между собой океан и материки.

Ограничимся здесь лишь последним видом потоков: потоками муссонного происхождения. Представим себе схематически, что в зимнее время в нижнем слое тропосферы толщиной в несколько сотен метров движется с некоторой средней скоростью холодный воздух по направлению с материка на океан; над ним в противоположном направлении — с океана на материк — движется с такой же средней скоростью более теплый воздух. Теория показывает, что количество тепла, вносимого муссонной системой воздушных потоков с океана на материк, пропорционально произведению средней скорости потока на разность средних температур верхнего и нижнего «этажей» муссонного слоя.

Сама скорость этого потока, в свою очередь, пропорциональна градиенту давления атмосферы, а последний связан с градиентом температуры тем основным уравнением муссонного поля, о котором была речь выше (в связи с анализом карт изобар и температурных изаномал).

Пусть та или иная причина нарушила установившееся движение воздушных потоков. Пусть, например, в некотором районе возникло местное потепление. Тогда, очевидно, по соседству с этим районом неминуемо наступит изменение градиентов температуры и градиентов атмосферного давления, связанных с ними тем же основным уравнением муссонного поля. Согласно известному принципу Лешателье — Брауна,

характер явлений должен при этом измениться в том направлении, чтобы вести к ликвидации возникшего возмущения в системе.

Следовательно, в районе, в котором возникло временное потепление, скорости воздушных потоков должны измениться таким образом, чтобы количество тепла, выделяющегося в нем, стало уменьшаться, ведя к ликвидации потепления. Как и всегда, во всякой системе, связанной с движением масс, обладающих инерцией, этот процесс не может ограничиться ликвидацией потепления: на смену потеплению в том же районе должно возникнуть местное похолодание. Но, очевидно, что после этого все рассмотренные процессы пойдут совершенно аналогично в обратном направлении, а это повлечет за собой новую перемену: на смену похолоданию опять в рассматриваемом районе наступит потепление.

Как видим, в атмосфере, пронизанной тепловыми потоками, местное нарушение теплового состояния неизбежно должно вести к нарушению распределения давлений, к изменению скоростей ветров. Следовательно, колебания температуры здесь должны быть неразрывно связаны с колебаниями атмосферного давления.

С другой стороны, никакое колебание температуры и никакое колебание давления в отдельном районе не может остаться изолированным: с ним непременно свяжутся подобные же колебания температур и колебания давлений в окружающих районах, подобно тому как нарушение покоя поверхности воды каким-нибудь брошенным камнем непременно вызовет вокруг потревоженного участка целые кольца волн.

В атмосфере, таким образом, должны возникать волнообразные колебания температур и волнообразные колебания давлений, передающиеся из одного района в другие. Ввиду неразрывности этих двух процессов назовем исследуемые волны термобарическими. Но всякие волны могут проявляться в двух несхожих между собой формах: либо они распространяются совершенно свободно без явно выраженного влияния каких-то границ, либо границы пространства, охваченного волнами, вызывают их отражение, а отраженные волны налагаются на основную систему волн, интерферируют с ними. Такое наложение приводит, как известно, к возникновению стоячих волн; на всем протяжении, охваченном ими, колебания приобретают совсем иной характер. На некоторых участках происходит одновременное повышение уровня,

а на иных участках одновременное падение уровня; участки первого и второго рода оказываются разделенными между собой некоторыми линиями, во всех точках которых уровень совершенно не меняется; это — так называемые узловые линии. Где-то в средней части каждого участка, заключенного внутри узловой линии, уровень колеблется сильней, чем во всех иных точках: здесь располагается пучность в системе стоячих волн.

В некоторых случаях стоячие волны могут создать очень сложный рисунок залегания узловых линий и пучностей, причем узловые линии будут проходить не только вдоль границ области, охваченной волнами, но и на протяжении самой области, где они образуют своеобразную сеть. Подобная сеть узловых линий возникает, например, на упругой пластинке, колебания которой возбуждаются смычком, проводимым по краю пластинки; мелкий песок, насыпанный на такую пластинку, собирается вдоль характерных линий, давая так называемые «хладниевы» фигуры. Не трудно заметить, что линии, отмеченные спокойно залегающими песчинками, представляют собой именно узловые линии на колеблющейся пластинке.

Совершенно аналогичные узловые линии возникают в системе термобарических волн, дробя поверхность земного шара на отдельные участки, в пределах которых проявляются самые типичные стоячие волны, или, как их еще называют, сейши. В настоящее время еще трудно установить окончательно, какие именно физико-географические условия определяют собой «рисунок» узловых линий и места расположения пучностей в системе термобарических сейш. Однако уже сейчас очевидно, что большое значение при этом должны иметь такие элементы земной поверхности, как горные цепи, берега океана и берега внутренних морей, протянувшихся друг за другом (как, например, берега Азовского, Черного и Средиземного морей).

Необходимо особо отметить, что всякие длинные волны, возникающие на поверхности Земли, должны испытывать на себе влияние вращения земного шара вокруг своей оси. Под действием этого вращательного движения Земли, как пучности термобарических сейш, так и узловые линии обязаны вращаться вокруг некоторых точек на земной поверхности, как это вытекает из теории относительного движения на вращающейся системе. Теория показывает также, что одновременно с подоб-

ным непрерывным вращением всей картины температура и давление атмосферы в пучностях должны пульсировать с тем же периодом, с каким происходит вращение этой картины.

Величина периода может быть определена теоретически по заданной разности температур воздуха между верхним и нижним «этажами» муссонного слоя атмосферы, т. е. между средней температурой воздуха, поступающего поверху на материк с океана, и средней температурой воздуха, стекающего понизу с материка на океан, по заданным размерам

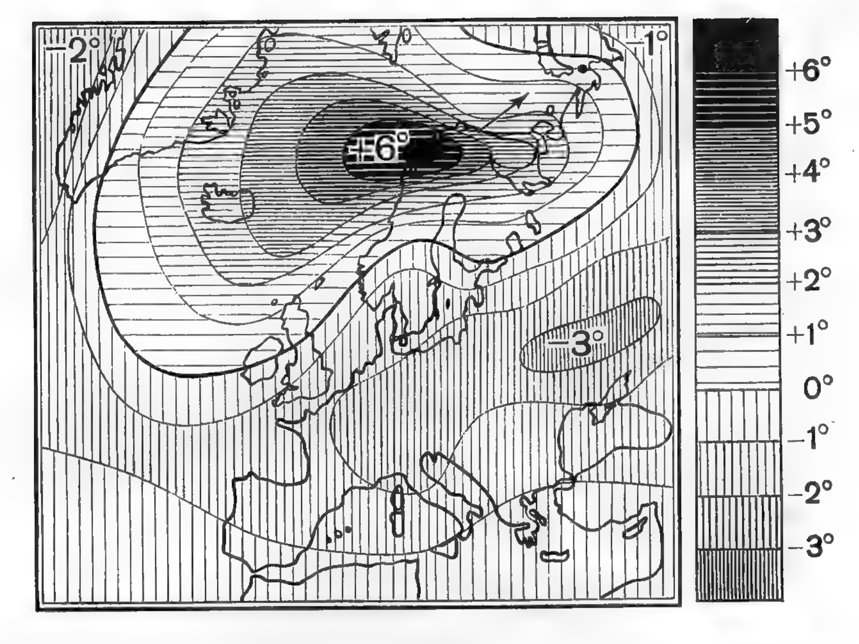


Фиг. 8

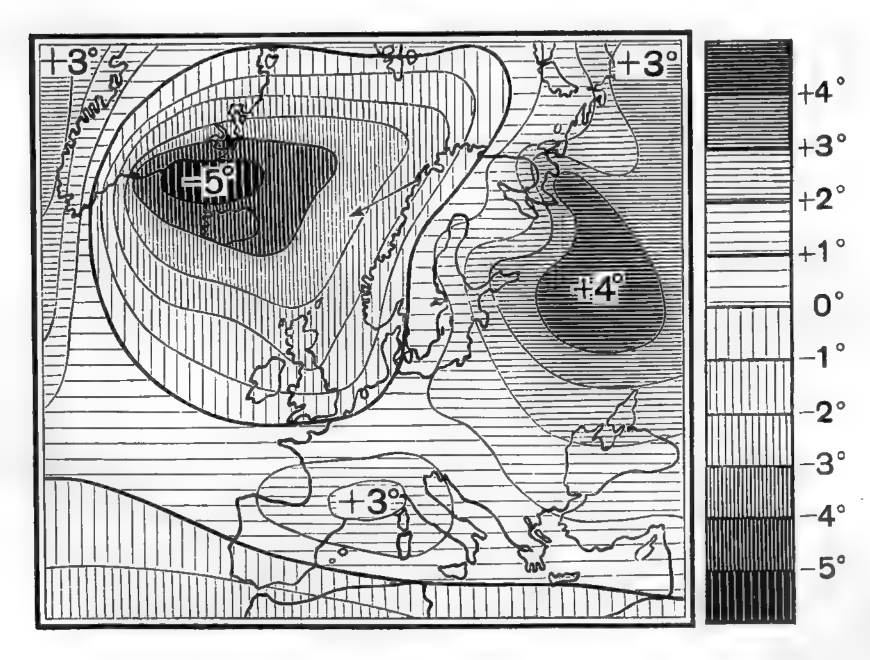
района, лежащего внутри узловой линии, и некоторым другим данным, на которых не будем здесь останавливаться.

Приближенное вычисление дало для периода наиболее часто проявляющихся термобарических сейш значение — 8 суток. Совершенно такое же числовое значение дают непосредственные наблюдения, как это видно, например, на фиг. 8, где воспроизведены кривые колебаний температуры воздуха и колебания атмосферного давления. Столь же замечательно, что отношение амплитуд колебания давления к амплитудам колебания температуры оказалось равным той же самой величине $\Pi = 1.6 \times 10^3$, которая была упомянута выше в связи с основным уравнением муссонного поля.

Но самое убедительное подтверждение нашей теории термобарических сейш получается при сопоставлении теоретически вычисленной картины их поля с картами отклонений температуры от среднемесячных



Фиг. 9



Фиг. 10

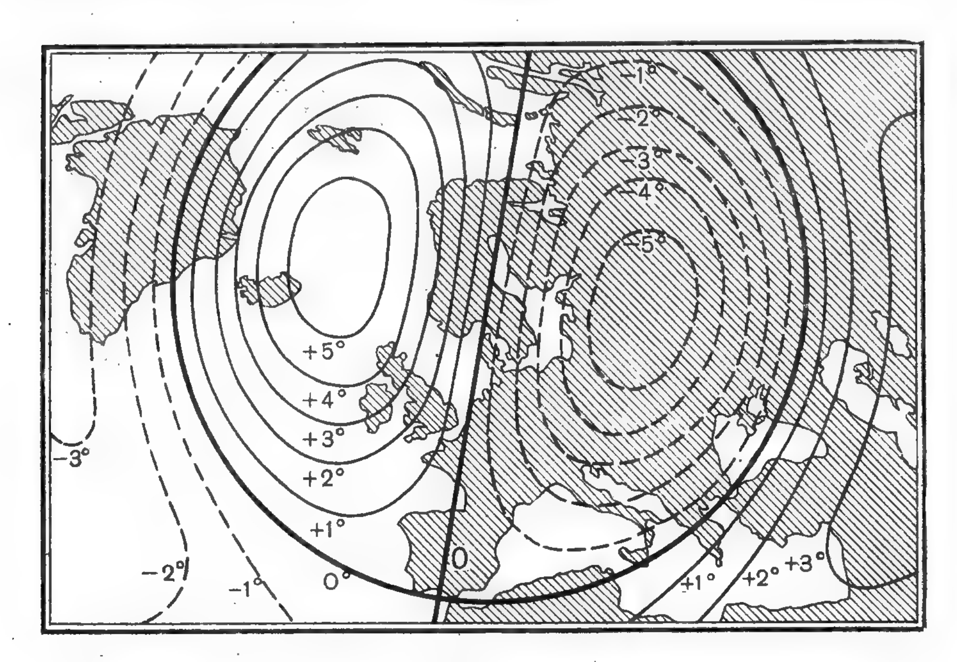
438 v 439

норм, которые были построены шведским геофизиком Сандстремом для Европы, Атлантического океана, Гренландии. Этот автор предполагал, что изменения направления ветра у Лофотенских островов свидетельствуют о каких-то пульсациях теплого Атлантического течения и Гольфстрима, служащего основой этого течения. Применительно к различным направлениям ветра у Лофотена (через каждые 10 градусов азимута) он построил осредненные за 80 лет карты потеплений и похолоданий воздуха в различных районах этой области. На фиг. 9 и 10 изображены две из его карт для двух противоположных фаз колебания температуры: для наибольшего потепления в Атлантике с одновременным наибольшим похолоданием в Европе и для наибольшего похолодания в Атлантике с одновременным наибольшим потеплением в Европе.

На фиг. 11 изображена наша теоретическая схема, построенная на основании анализа термобарических сейш в атмосфере. Как видно, она очень хорошо согласуется с действительностью: на ней явственно видны те же пучности температурных колебаний и тот же характер узловых линий, как и на фиг. 9 и 10.

Теория показывает, что термобарические сейши могут возникать также и в системе тепловых потоков, направленных вдоль меридиана и обусловленных неодинаковым прогревом различных широтных зон. В некоторых, довольно редких, случаях колебания в меридиональных потоках могут попасть в резонанс с колебаниями в потоках муссонного происхождения. Последствия подобного резонанса, по всей вероятности, проявились во время знаменитых морозов зимой 1940 года. Вопреки шаблонной гипотезе о «вторжении полярных масс», посредством которой современные синоптики склонны объяснять все резкие похолодания, морозы 1940 года ударили прежде всего по югу нашей страны одновременно с оттепелью, установившейся на севере; только вслед за этим на юге наступила оттепель, а на севере — небывалые морозы, особо суровые по соседству с Ладожским озером. С точки зрения нашей теории, холодные и теплые массы находятся в состоянии непрерывной «борьбы»; в самом начале описываемого промежутка времени тепловые потоки, идущие вдоль меридиана, оставили меньше тепла в южной части нашей европейской территории и пронесли его дальше на север (по сравнению с нормой); это повело к временному сильному похолоданию на Украине и в Крыму; вслед за тем, на Украине и в Крыму на подогревание

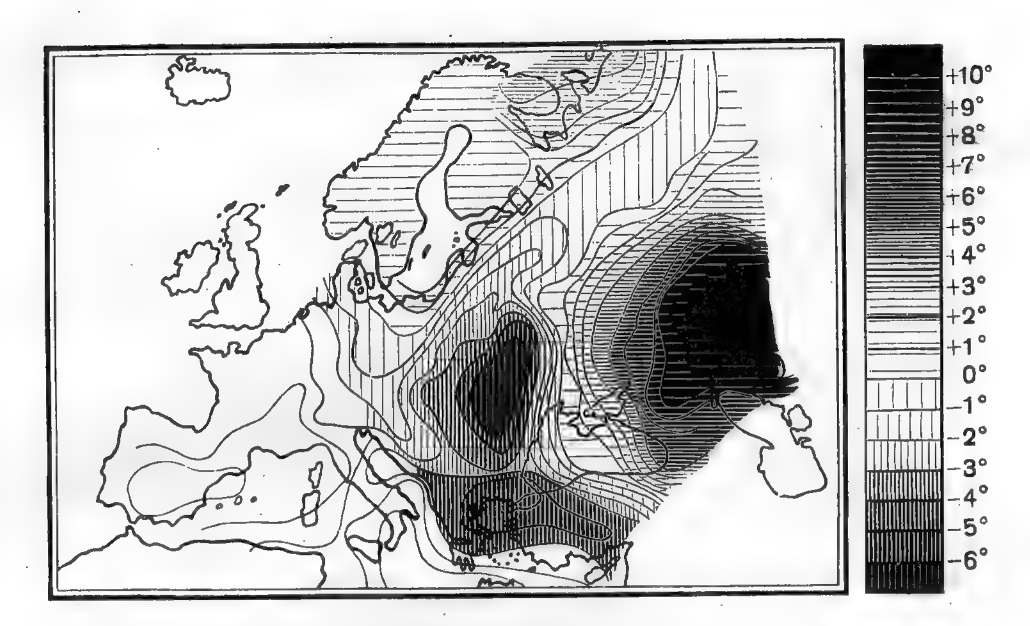
воздуха было израсходовано повышенное количество тепла, шедшего вдоль меридиана, а потому в северной части нашей европейской территории возник недогрев воздуха, возникли суровые морозы. По всей вероятности, одновременно с колебаниями в системе меридиональных



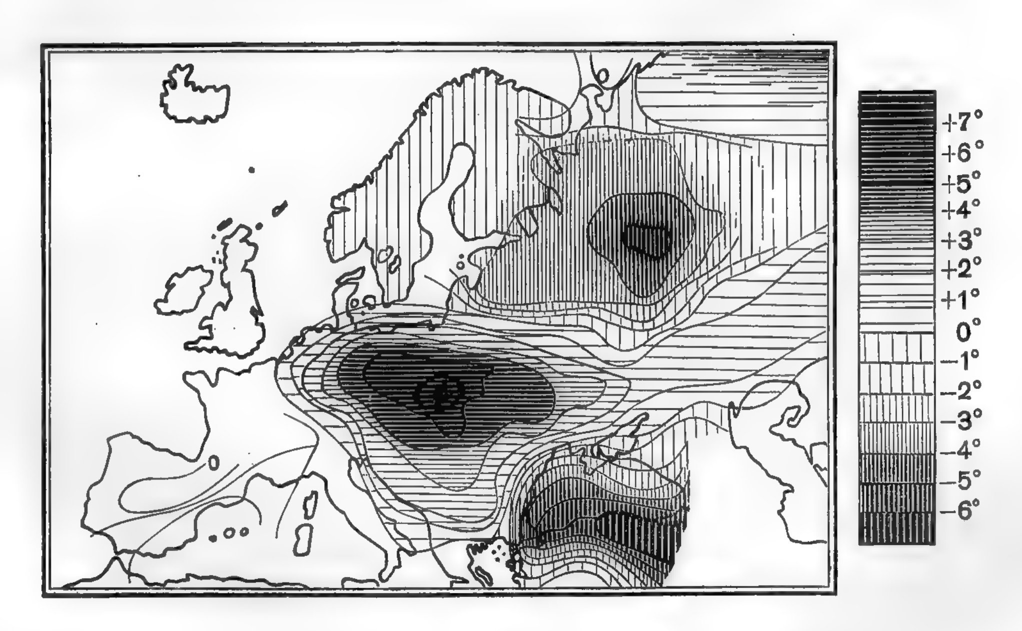
Фиг. 11

потоков тепла здесь происходили колебания в муссонных потоках, которые необычайно обострили явление.

До настоящего времени можно лишь догадываться о том, что термобарические сейши возникают в атмосфере преимущественно за счет тех возмущений, которые связаны с возникновением громадных вихрей — циклонов на поверхностях раздела холодных и теплых воздушных масс. Мы еще ничего не знаем ни о деталях этого процесса зарождения колебаний, ни о причинах, способствующих или их пышному развитию, или их быстрому затуханию. Можно лишь догадываться о том, что степень развития термобарических сейш должна находиться в прямой связи



Фиг. 12



Фиг. 13

440 v 441

с отношением двух периодов: периода собственных колебаний системы внутри той или иной узловой линии и промежутков времени между отшнуровыванием последовательных вихрей — циклонов в районе их преимущественного зарождения (над Атлантическим океаном).

Но если подобные колебания постоянно возникают в атмосфере на всем протяжении года, то естественно ожидать их особого обострения во время с м е н ы м у с с о н н ы х с е з о н о в — весной и осенью. В самом деле, ведь именно весной и осенью достигает наибольших скоростей переход дополнительных воздушных масс с океана на материк и в обратном направлении. И действительно, именно весной и осенью разыгрываются в атмосфере наиболее мощные термобарические сейши: они давным давно наблюдались в форме так называемых «майских холодов» и так называемого «бабьего лета», но только в настоящее время они стали совершенно понятными с точки зрения теории.

На фиг. 12 и 13 изображены карты потеплений и похолоданий в Европе в сезон майских холодов (заимствованные из студенческой дипломной работы, проделанной на Черноморской гидрофизической станции АН СССР). На них отражены две фазы колебательного процесса, протекающего по той же описанной схеме: пучности и узловые линии вращаются против часовой стрелки вокруг некоторой точки (в данном случае точка эта лежит недалеко от г. Пензы, там же, где лежала она во время упомянутых морозов 1940 года) с периодом, значительно превышающим 8 суток, — с периодом около 26 суток. Такое увеличение периода объясняется рядом естественных причин: прежде всего, колебания охватывают здесь значительно больший район; разность температур между нижним и верхним «этажами» муссонного слоя здесь меньше, чем бывает зимой, а это должно отразиться на увеличении периода; наконец, в отличие от обычной картины (фиг. 9—11) здесь (фиг. 12 и 13) видны ч е т ы р е пучности температурных колебаний.

За время одного оборота такого «четырехполюсного колеса», очевидно, должны четыре раза возникнуть смены: от похолодания к потеплению и в обратном порядке.

Так как потепление в весеннее время всегда представляется естественным, то все внимание привлекают те фазы колебательного процесса, которые соответствуют похолоданию; именно это похолодание хорошо известно в сельском хозяйстве, где оно нередко причиняет большие беды.

Совершенно аналогичная картина, но в обратном направлении, наблюдается осенью: на фоне общего естественного похолодания возникают кратковременные возвраты тепла, которые также известны с давних пор и которые получили название «бабьего лета».

Как видим, теория термобарических сейш позволяет рационально объяснить процесс смен погоды, не прибегая к гипотезе о самопроизвольно скитающихся «холодных массах» и «теплых массах». В действительности все колебания среднего режима грандиозных «тепловых машин» в атмосфере проявляются в виде естественного перераспределения одного и того же теплового бюджета воздушных потоков, перераспределения, которое попеременно приводит к потеплению в одном районе за счет похолодания в другом и к последующему похолоданию в первом районе за счет потепления во втором.

Но ведь легко себе представить подобное же перераспределение теплового бюджета, установленного Солнцем, в более широком масштабе: перераспределение может охватить не только мощные слои атмосферы, до стратосферы включительно (а совсем не один «муссонный» слой, достигающий всего лишь 2 км в высоту), но и еще более теплоемкие слои океанических вод. Совершенно очевидно, что при таких коренных перераспределениях солнечного теплового бюджета должны возникать значительные изменения самого климата в различных областях материков и океанов. Именно здесь, в недрах океана и на высотах атмосферы, следует искать разгадку перемен *климата* Земли при практически неменяющейся солнечной деятельности, которую характеризует вполне оправдывающая свое название «солнечная постоянная». Было бы явно нелепым искать в космических влияниях причину тех резких изменений климата, которые разыгрываются в относительно короткие сроки на наших глазах. Мы *обязаны* до конца использовать изучение тех изменений теплового режима, которые происходят в воздухе, над кровлями наших городов, и в океане, под килями наших кораблей.

Член-корреспондент АН СССР
И. А. Кибель

ПРОГНОЗ ПОГОДЫ КАК ЗАДАЧА ДИНАМИЧЕСКОЙ МЕТЕОРОЛОГИИ

1. Советская школа динамической метеорологии, к которой я принадлежу и о работах которой я буду здесь говорить, была основана проф. А. А. Фридманом. А. А. Фридман широко известен среди специалистов по гидромеханике как один из творцов статистической теории турбулентности; физики знают А. А. Фридмана как ученого, вышедшего победителем из полемики с Альбертом Эйнштейном. Но для нас, советских метеорологов, А. А. Фридман прежде всего создатель нового направления в нашей науке. Для этого направления было лозунгом — бросить на службу советской метеорологии все современные достижения физики, в особенности гидродинамики и термодинамики.

Н. Е. Кочин был одним из первых представителей этого направления, а после смерти А. А. Фридмана (1925) возглавил молодую школу. К этой школе примыкают П. Я. Полубаринова, Л. В. Келлер, А. А. Дородницын, Е. Н. Блинова, М. И. Юдин, М. Е. Швец, Н. Р. Малкин и др.

В результате усилий школы Фридмана — Кочина динамическая метеорология стала непосредственно, без промежуточных инстанций, применяться при прогнозе погоды.

2. В 1913 г. в Лейпциге при вступлении в должность профессора В. Бьеркнесс произнес речь, которую он озаглавил «Die Meteorologie als exakte Wissenschaft». В ней он говорил о том, что система уравнений

гидродинамики вместе с некоторыми законами термодинамики должна служить для решения метеорологических задач. Он не предлагал никакого метода решения этой системы и лишь высказывал предположение, что работа окажется весьма трудоемкой и, может быть, потребуется создать специальные методы расчета (графический и т. п.).

В 1921 г. в замечательной книге «Weather Predication by Numerical Process» Л. Ричардсон уже конкретно подошел к вопросу. Основная его идея очень проста: уравнения гидродинамики вместе с первым законом термодинамики представляют пять уравнений с пятью неизвестными функциями (компоненты скоростей, давление, температура). Эти уравнения можно решить относительно первых производных по времени от пяти перечисленных функций, так что в правые части будут входить лишь производные по координатам. Дальше остается техническая работа — дифференцирование и составление рядов Тэйлора. К сожалению, вычисление конкретного примера прогноза, проведенное Ричардсоном, заняло очень много времени, да и сам прогноз оказался неверным. Причина неудачи была связана с необычайной громоздкостью его формул (он хотел охватить огромное число объектов сразу — испарение с листьев, температуру почвы и т. д.); кроме того, он мог оперировать лишь с приземными данными (радиозонды появились лет на 15 позже), и начальные данные были, естественно, неполны, да и нерепрезентативны. Гидродинамика того времени была не подготовлена к решению подобных задач; не было начальных данных для верхних слоев. Через 20 лет картина существенно изменилась.

3. Сейчас, когда мы говорим о прогнозе метеорологических элементов при помощи решения уравнений физики, мы начинаем с того, что представляем себе атмосферу как воздушную оболочку, вращающуюся, по крайней мере в интересующих прогнозиста слоях, вместе с Землей, т. е. подверженную действию силы тяжести и силы Кориолиса. Мы говорим затем о притоке тепла к атмосфере — о солнечной и длинноволновой радиациях, постоянное наличие которых обусловливает, в частности, те распределения температуры по высоте, которые мы приписываем так называемой стандартной атмосфере. Мы вспоминаем о роли подстилающей поверхности и постоянно существующей атмосферной турбулентности. Мы выделяем большую группу вопросов, связанных с распределением влажности, ее конденсацией, выпадением.



Эти факторы, постоянно действующие, создают, так сказать, фон явления. В задаче собственно прогноза, т. е. в задаче о том, каково будет далее распределение давления, температуры, ветра, влажности, если в начальный момент эти элементы заданным образом распределены в пространстве, фон этот играет своеобразную роль. С одной стороны, мы не имеем права говорить о прогнозе (на короткий или на долгий срок) при помощи решения уравнений физики, если из этих же уравнений, может быть иначе обработанных, мы не можем получить распределение метеорологических элементов фона. С другой стороны, напротив, мы можем с самого начала поставить вопрос о существующих в начальный момент отклонениях от стандартных значений и искать поведение этих отклонений далее. Эти две постановки задачи постоянно переплетаются друг с другом.

4. Уравнения движения воздуха, если иметь в виду краткосрочный прогноз погоды, принято в метеорологии писать в виде

$$\frac{du}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + 2\omega_z v + R_x, \tag{1}$$

$$\frac{dv}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial y} - 2\omega_z u + R_y, \tag{2}$$

$$\left(\frac{d}{dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + w\frac{\partial}{\partial z}\right),$$

$$0 = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g. \tag{3}$$

Здесь x и y — горизонтальные координаты, z — высота над уровнем моря, t — время, u, v — составляющие скоростей по осям x, y соответственно, $\omega_z = \omega \sin\varphi$, где $\omega = 7.29 \cdot 10^{-5}$ сек$^{-1}$ — угловая скорость вращения земли, φ — широта места,[1] p — давление, ρ — плотность, g — ускорение силы тяжести. Речь идет, конечно, о средних, сглаженных значениях наших элементов, поэтому в уравнение вошли члены R_x, R_y, изображающие турбулентное трение. В нашей задаче мы принимаем

[1] Широта той точки, в которой помещается начало координат.

R_x, R_y в виде

$$R_x = \frac{\partial}{\partial z}\left(\frac{\mu}{\rho}\frac{\partial u}{\partial z}\right) + \frac{\mu'}{\rho}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right),$$

$$R_y = \frac{\partial}{\partial z}\left(\frac{\mu}{\rho}\frac{\partial v}{\partial z}\right) + \frac{\mu'}{\rho}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right), \quad (4)$$

где μ, функция от z, — коэффициент турбулентной вязкости для вертикального перемешивания, а μ' — принимаемый независящим от высоты коэффициент горизонтального перемешивания. Уравнение неразрывности имеет вид

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} + \frac{\partial \rho v}{\partial z} = 0. \quad (5)$$

Уравнение притока тепла (первый закон термодинамики) мы должны написать в виде

$$c_p \rho \frac{dT}{dt} - \frac{1}{J}\frac{dp}{dt} = \varepsilon_1 + \varepsilon_2, \quad (6)$$

где T — температура, c_p — теплоемкость при постоянном давлении, J — механический эквивалент тепла, ε_1 и ε_2 — два вида притока тепла к частице: от турбулентной теплопроводности и от радиации.[1]

Мы принимаем обычно, что ε_1 имеет вид

$$\varepsilon_1 = \frac{\partial}{\partial z}\left(\lambda \frac{\partial T}{\partial z}\right) + \lambda'\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right), \quad (7)$$

где λ, λ' — коэффициенты турбулентной теплопроводности, и что ε_2 можно представить по Шварцшильду Эмдену в виде

$$\varepsilon_2 = \alpha \rho_w (A + B + \beta S - 2E). \quad (8)$$

Здесь A и B — потоки длинноволновой радиации, идущие сверху и снизу соответственно, S — поток солнечной радиации,

$$E = f\sigma T^4, \quad (9)$$

где σ — постоянная Стефана-Больцмана, f — правильная дробь, множитель, введенный Гульбертом, ρ_w — плотность поглощающего вещества

[1] Мы не рассматриваем на первых порах теплоту, которая в тех или иных случаях освобождается за счет конденсации паров воды или идет на испарение капелек.

(в основном паров воды), α — коэффициент поглощения для «серой» длинноволновой радиации, β — отношение коэффициентов поглощения коротковолновой и длинноволновой радиаций.

Если считать μ, λ, μ', λ', $\alpha\rho_w$ известными, то и тогда задача сводится к определению 9 функций u, v, w, p, ρ, T, A, B, S из пяти уравнений (1), (2), (3), (5), (6). Недостающие уравнения не трудно построить. Это, во-первых, закон Клапейрона

$$p = R\rho T, \tag{10}$$

где R — газовая постоянная; и, во-вторых, это законы, регулирующие поглощение и испускание излучения среды

$$\frac{\partial A}{\partial z} = \alpha\rho_w (A - E), \tag{11}$$

$$\frac{\partial B}{\partial z} = \alpha\rho_w (E - B), \tag{12}$$

$$\frac{\partial S}{\partial z} = \alpha\rho_w S. \tag{13}$$

Прежде чем говорить о прогнозе, прежде чем задавать начальные условия, посмотрим, каковы краевые условия нашей задачи и как получается тот фон, о котором мы упоминали выше.

На поверхности земли, точнее на «уровне шероховатости», естественно принять $u = v = w = 0$. Коэффициент μ сперва растет с высотой, начиная от малых (близких к величинам лабораторного коэффициента вязкости воздуха) значений, вплоть до величин порядка 50 г/см сек.; эта величина достигается на уровне, примерно, в 50—100 м, причем далее μ меняется значительно меньше. Коэффициенту μ' принято придавать разные значения в зависимости от масштабов изучаемого явления; так, при построении теории климата принимают $\mu' \approx 10^6\ \mu$; в задачах краткосрочного прогноза $\mu' \approx 10^2\ \mu$. Если принять в расчет, что в обычных условиях величина $\frac{\partial^2 u}{\partial x^2}$ $\left(\text{или } \frac{\partial^2 u}{\partial y^2}\right)$ будет примерно в 10^6 раз меньше, нежели $\frac{\partial^2 u}{\partial z^2}$, то станет ясно, что в задаче о краткосрочном прогнозе можно записать просто

$$R_x \approx \frac{\partial}{\partial z}\left(\frac{\mu}{\rho}\frac{\partial u}{\partial z}\right), \qquad R_y \approx \frac{\partial}{\partial z}\left(\frac{\mu}{\rho}\frac{\partial v}{\partial z}\right).$$

Наибольшие изменения величины $\frac{\partial u}{\partial z}$ $\left(\text{или } \frac{\partial v}{\partial z}\right)$ следует ожидать благодаря краевым условиям у поверхности земли. Если δ — характерная высота, до которой происходит быстрое изменение градиентов, то, сравнивая порядок членов типа $2\omega_z v$ и R_x, придем к выводу, что «толщина» δ должна быть

$$\delta \approx \sqrt{\frac{\mu_0}{2\rho_0\omega_z}},$$

где μ_0, ρ_0 — характерная турбулентная вязкость и характерная плотность соответственно. Это так называемый планетарный пограничный слой атмосферы, всегда существующий слой, в котором турбулентность проявляется. Толщина его для наших широт будет иметь порядок 500 м. Выше этого слоя, вообще говоря, турбулентной вязкостью можно пренебречь, и мы можем поставить задачу так, как она ставится в обычной теории пограничного слоя в гидродинамике: сперва отбросить полностью члены с R_x и R_y и решать задачу с краевыми условиями, отвечающими идеальной жидкости (на поверхности $w=0$), а затем сделать поправку на пограничный слой.

Введем обозначения

$$G_x = -\frac{1}{2\omega_z\rho}\frac{\partial p}{\partial y}, \quad G_y = \frac{1}{2\omega_z\rho}\frac{\partial p}{\partial x}. \tag{14}$$

Введем характерное время t_0, характерную длину l и характерную горизонтальную скорость $V\left(=\frac{l}{t_0}\right)$. Обозначим

$$t = t_0\bar{t}, \quad x = l\bar{x}, \quad y = l\bar{y}, \quad u = V\bar{u}, \quad v = V\bar{v}.$$

Уравнения (1), (2) вне пограничного слоя можно записать теперь в виде

$$\frac{1}{2\omega_z t_0}\left(\frac{\partial}{\partial\bar{t}} + \bar{u}\frac{\partial}{\partial\bar{x}} + \bar{v}\frac{\partial}{\partial\bar{y}} + \bar{w}\frac{\partial}{\partial\bar{z}}\right)u = -G_y + v,$$

$$\frac{1}{2\omega_z t_0}\left(\frac{\partial}{\partial\bar{t}} + \bar{u}\frac{\partial}{\partial\bar{x}} + \bar{v}\frac{\partial}{\partial\bar{y}} + \bar{w}\frac{\partial}{\partial\bar{z}}\right)v = G_x - u. \tag{15}$$

Пусть теперь характерное время t_0 будет величиной порядка суток и более. Тогда

$$\varepsilon = \frac{1}{2\omega_z t_0} \tag{16}$$



Н Кочин

448 и 449

будет иметь порядок 1/10 и меньше, т. е. наши уравнения (15) будут содержать малый параметр при старших производных u и v. Грубо говоря, это значит, что всегда будет просто

$$u \approx G_x, \qquad v = G_y.$$

Это так называемый геострофический ветер.

Точный анализ показывает, что решения представляются в виде рядов по степеням ε вида

$$u = G_x + u' + \ldots, \qquad v = G_y + v' + \ldots \tag{17}$$

Члены с ε в первой степени суть u', v'; мы дадим их выражения ниже, а сейчас обратимся к уравнению неразрывности (5).

Порядок величины $\frac{\partial \rho}{\partial t}$ будет 10^{-10} г/см^3 сек., в то время как порядок $\frac{\partial \rho w}{\partial z}$ будет 10^{-9} г/см^3 сек., а порядок $\frac{\partial \rho u}{\partial x}$ $\left(\text{или } \frac{\partial \rho v}{\partial y}\right)$ — даже 10^{-8} г/см^3 сек. Но в силу (14) и (17) мы можем написать

$$\frac{\partial \rho u}{\partial x} + \frac{\partial \rho v}{\partial y} = \frac{\partial \rho u'}{\partial x} + \frac{\partial \rho v'}{\partial y},$$

так что

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u'}{\partial x} + \frac{\partial \rho v'}{\partial y} + \frac{\partial \rho w}{\partial z} = 0.$$

Три последних члена в этом уравнении имеют теперь один и тот же порядок малости — 10^{-9} г/см^3 сек.; первый будет по крайней мере в 10 раз меньше. Выделение в указанных трех членах новых главных членов в общем виде представляет огромные трудности. Большие усилия иностранных ученых, направленные к этой цели, до сих пор не дали существенных результатов. Мы ограничимся тем, что положим просто

$$\frac{\partial \rho w}{\partial z} + \frac{\partial \rho u'}{\partial x} + \frac{\partial \rho v'}{\partial y} = 0$$

и для ρw примем выражение

$$\rho w = -\int_0^z \left(\frac{\partial \rho u'}{\partial x} + \frac{\partial \rho v'}{\partial y}\right) dz \tag{18}$$

(земля считается плоской). Заметим попутно, что w, представляясь через u', v', уже само будет содержать наш множитель ε.

Поэтому окончательно

$$u' = -\frac{1}{2\omega_z}\left(\frac{\partial}{\partial t} + G_x\frac{\partial}{\partial x} + G_y\frac{\partial}{\partial y}\right)G_y,$$

$$v' = \frac{1}{2\omega_z}\left(\frac{\partial}{\partial t} + G_x\frac{\partial}{\partial x} + G_y\frac{\partial}{\partial y}\right)G_x. \tag{19}$$

Формулы (17) — (19) позволяют найти составляющие скоростей, как только известно распределение p и ρ. Закон Клапейрона позволит исключить ρ. Барометрическая формула — следствие уравнения (3)

$$p(x, y, z, t) = p_0(x, y, t)\, e^{-\frac{g}{R}\int\limits_0^z \frac{dz}{T(x, y, z, t)}} \tag{20}$$

представит p через наземное давление p_0 и через температуру T. Задача сводится к отысканию двух функций p_0 и T. Для определения этих функций будут служить уравнение (6) и условие на бесконечности.

Все члены левой части уравнения (6) имеют примерно один и тот же порядок, что хорошо будет видно, если ввести наши безразмерные величины и написать

$$c_p\rho\frac{dT}{dt} - \frac{1}{J}\frac{dp}{dt} =$$

$$= \frac{c_p\rho}{t_0}\left\{\left(\frac{\partial}{\partial \bar{t}} + \bar{u}\frac{\partial}{\partial \bar{x}} + \ldots\right)T - \frac{\varkappa - 1}{\varkappa}\frac{T}{p}\left(\frac{\partial}{\partial \bar{t}} + \bar{u}\frac{\partial}{\partial \bar{x}} + \ldots\right)p\right\}$$

($\varkappa$ — отношение теплоемкостей). Эти члены имеют порядок 10^{-8} (CGS). Напротив, члены, входящие в ε_2, согласно (8), имеют все порядок 10^{-7}; тот же порядок имеет ε_1 (там, где турбулентная теплопроводность проявляется, т. е. в местах с большим изменением вертикальных температурных градиентов). Было бы, однако, непростительной ошибкой, если бы мы на основании этого отбросили левую часть (6) в нашей задаче о прогнозе. Коренное отличие уравнения (6) от уравнений движения и уравнения неразрывности в том, что правая часть (6) содержит постоянно действующие в атмосфере факторы (A, B, E, S, λ). Отбросив левую часть (6) и исключив при помощи (11) и (12) A и B, мы получим для определения T

(или E) уравнение, из которого мы просто найдем стандартное распределение температуры (атмосфера с ее тропосферой и характерным переходом в стратосферу).

Чтобы выяснить, как же происходит движение, сохраним в левой части (6) все члены и введем вместо высоты z приведенную оптическую толщину (для длинноволновой радиации) из равенства

$$\zeta = \frac{1}{\tau_0}\int_z^\infty \alpha\rho_w dz, \text{ где } \tau_0 = \int_0^\infty \alpha\rho_w dz. \qquad (21)$$

Деля левую часть (6) на $\alpha\rho_w$ и вводя вместо температуры функцию E из (9), получим сперва для левой части

$$\frac{1}{\alpha\rho_w}\left(c_p\rho\frac{dT}{dt} - \frac{1}{J}\frac{dp}{at}\right) = N\left\{\frac{\partial E}{\partial \bar{t}} + \ldots\right\}, \qquad (22)$$

где безразмерный параметр N имеет вид

$$N = \frac{c_p\rho}{4x\rho_w f\sigma T^3 t_0} \qquad (23)$$

[в (22) выписан явно лишь первый член]. Исключая затем A и B из (6) при помощи (11) и (12), получим следующее дифференциальное уравнение 4-го порядка

$$\frac{2}{m^2-1}\left(\frac{1}{\tau_0^4}\frac{\partial^4}{\partial z^4} - \frac{m^2}{\tau_0^2}\frac{\partial^2}{\partial z^2}\right)E - N\left(\frac{1}{\tau_0^2}\frac{\partial^2}{\partial z^2} + 1\right)\left\{\frac{\partial E}{\partial \bar{t}} + \ldots\right\} = \beta(1-\beta^2)S, \qquad (24)$$

где m — новый безразмерный параметр такой, что

$$\frac{m^2-1}{2} = \frac{\alpha\rho_w\lambda}{4f\sigma T^3}.\text{[1]} \qquad (25)$$

Величина $\frac{m^2-1}{2}$ имеет порядок 1.[2] Величина N, если за характерное время примем сутки, будет также

$$N \approx 1.$$

---

[1] Приближенно можно принять m и N постоянными, взяв их средние значения; это и делается при выводе (24).

[2] При подсчете мы принимаем сейчас и далее у земли $\alpha\rho_w \approx 4\cdot 10^{-5}$ см$^{-1}$.

Оптическая толщина же атмосферы τ_0 имеет порядок

$$\tau_0 \approx 13.$$

Таким образом, наши уравнения содержат малые коэффициенты при старших производных по ζ. Это — характерное свойство уравнений, описывающих пограничные слои. Для таких уравнений краевые условия действуют лишь в очень малой области, примыкающей к границе. Наше уравнение — 4-го порядка по ζ, и мы должны для него написать четыре краевых условия. Два из этих условий берутся на верхней границе атмосферы ($z=0$, т. е. $\zeta=0$). Это — условие отсутствия там A (потока длинноволновой радиации, идущей сверху) и условие затухания изменений E. Два других условия записываются на поверхности земли. Это, во-первых, условие «баланса»

$$-\lambda \frac{\partial T}{\partial z} + \lambda^* \frac{\partial T^*}{\partial z} = A + S - B, \quad \text{при} \quad z=0, \tag{26}$$

где λ^*, T^* — теплопроводность и температура подстилающей поверхности, и, во-вторых, условие того, что B (радиация идущих вверх потоков) у самой земли создается именно подстилающей поверхностью

$$B = q\sigma T^4,$$

где q близко к единице.

Из-за такого разбиения краевых условий на две пары наше уравнение обладает д в у м я «пограничными слоями». Один из них примыкает к земле и занимает «расстояние» от $\zeta=1$ до $\zeta=0.9$ приблизительно; другой идет вниз от $\zeta=0$ до $\zeta=0.1$. Первый из этих слоев примерно совпадает с планетарным пограничным слоем, о котором мы говорили выше, когда приводили данные о ветре. Второй захватывает в с ю с т р а т о с ф е р у. Вне этих пограничных слоев, т. е. во всей тропосфере, за исключением планетарного пограничного слоя, мы можем написать

$$N\left\{\frac{\partial E}{\partial t} + \ldots\right\} = \beta(1-\beta^2)S,$$

или, так как правая часть из-за наличия β очень мала, написать, возвращаясь к прежним обозначениям,

$$c_p \rho \frac{dT}{dt} - \frac{\varkappa - 1}{\varkappa} \frac{T}{p} \frac{dp}{dt} = 0. \tag{27}$$

Это — условие адиабатичности, и им мы можем пользоваться в указанной области, если характерное время t_0 будем иметь порядка суток. Но, пользуясь этим уравнением, мы должны заменить то граничное условие затухания на бесконечности, которое мы поставили, когда решали уравнение 4-го порядка, некоторым условием на верхней границе интересующей нас области, т. е. условием на тропопаузе. В 1940 г. я принял в качестве такого условия описание того факта, что тропопауза состоит из одних и тех же частиц. Приближенно это приводит к соотношению

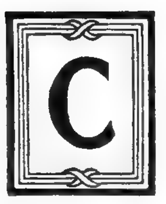
$$\frac{\partial T}{\partial t} + u_H \frac{\partial T}{\partial x} + v_H \frac{\partial T}{\partial y} = 0 \tag{28}$$

(на уровне тропопаузы), где u_H, v_H — составляющие скоростей на тропопаузе. Уравнения (27) и (28) вместе с барометрической формулой (20) и представлениями для скоростей и позволили мне дать способ решения задачи о прогнозе давления, температуры и ветра на срок порядка суток вперед. По исключении из (27) и (28) скоростей и введении p_0 и T в качестве искомых функций, мы приходим к уравнениям, содержащим слева $\frac{\partial p_0}{\partial t}$, $\frac{\partial T}{\partial t}$, а справа — производные от p_0, T по координатам. Дальше уже можно итти хотя бы по Ричардсону, отыскивая все новые производные по времени и представляя в каждой точке решение в виде ряда Тэйлора по t. Расчет по этим формулам сведен в настоящее время к очень простой схеме и отнимает менее получаса для прогноза давления, температуры и ветра в одном пункте на 24—36 час. вперед. Методика эта употребляется в практической работе некоторых институтов Гидрометеорологической службы СССР.

Работа, о которой я говорил, была продолжена и расширена в исследованиях других представителей школы Фридмана-Кочина. Труды А. А. Дородницына, М. Е. Швеца, Е. Н. Блиновой, опубликованные в разное время в периодических изданиях Академии Наук СССР, стоят здесь на первом месте.

Профессор
В. Ф. Бончковский

СЕЙСМИЧЕСКОЕ РАЙОНИРОВАНИЕ ТЕРРИТОРИИ СССР И ЗАДАЧИ АНТИСЕЙСМИЧЕСКОГО СТРОИТЕЛЬСТВА



Сейсмическое районирование территории СССР — одна из важнейших проблем в планировании строительства в сейсмически неспокойных районах. Нормы строительства, выработанные для землетрясений различной силы, имеют применение именно на основе карты сейсмического районирования и служат к сохранению не только материальных ценностей, но и жизни людей во время неожиданных катастрофических землетрясений. За время советской власти в Академии Наук СССР была проделана большая созидательная работа, которой и посвящен настоящий краткий очерк.

Многие участки территории СССР подвержены действию землетрясений. Юго-западные районы Украины нередко испытывают колебания почвы, источником которых служат сейсмические очаги в Карпатах. При катастрофических землетрясениях в Карпатах ощущаются колебания в Москве и даже Ленинграде, как было, например, отмечено в 1802 и 1940 гг.

Крымский полуостров непосредственно примыкает к довольно активной группе очагов, лежащих близ его южного берега. Списки землетрясений Крыма содержат описания почти 100 случаев, не считая нескольких сот повторных толчков, последовавших за сильными землетрясениями 1927 г.

На Кавказе ежемесячно регистрируется несколько десятков землетрясений. Исторические хроники Армении, Азербайджана, Грузии содержат известия о грандиозных сейсмических катастрофах, повлекших за собой многие тысячи человеческих жертв и разрушение городов.

Из 130 ощутимых землетрясений в Туркмении некоторые достигали катастрофической силы: Красноводское в 1895 г., Ашхабадское в 1929 г., последнее Казанджикское в 1946 г. Повторные толчки в каждом таком случае насчитываются сотнями. Очаги связаны с южными горными районами.

Восточные части Средней Азии — территории Узбекистана, Таджикистана, Киргизии и юго-восточные районы Казахстана — оказываются средоточием основной массы землетрясений, известных в СССР. Сейсмическими станциями здесь отмечается до тысячи землетрясений в год; по силе многие землетрясения достигали 10—12 баллов, т. е. предельной интенсивности (Верненское 1887 г., Каратагское 1907 г., Кебинское 1911 г., последнее Нарынское 1946 г. и многие другие). Особенно большой частотой толчков отличаются центральные и восточные районы Таджикистана.

Далее к северо-востоку землетрясения известны всюду в пределах горных районов — Алтай, Кузбасс, Саяны. Очень активна в сейсмическом отношении Байкальская впадина и аналогичные ей депрессионные формы Прибайкалья — долины Тунки, Иркута и т. д.

Эпизодические случаи известны и восточнее, вдоль железной дороги, вплоть до берегов Дальневосточного края. Множество сейсмических очагов встречается в полосе Курильских островов, в южной и восточной частях Камчатки. Известны землетрясения и во всей Восточной Сибири — между Леной и Анадырским краем.

Наконец, четкая полоса эпицентров протягивается от Верхоянского хребта, по островам арктических морей к Шпицбергену, где сливается с атлантической зоной эпицентров.

Отдельные случаи колебаний почвы фиксируются изредка даже на Урале, в Европейской части СССР, в том числе на Кольском полуострове.

За последние 15 лет на территории СССР отмечено не менее 5000 значительных землетрясений, таких, которые ощущались непосредственно людьми и занесены в каталоги. Кроме того, каждое крупное землетрясение сопровождается множеством повторных толчков. Сейсмические станции отмечают в год около 2000 землетрясений, эпицентры которых

ложатся в пределах СССР. Большинство землетрясений относится к слабым, нередко встречаются сильные и в единичных случаях — катастрофические.

Отсюда понятен интерес, который всегда возбуждали в нашей стране сейсмические явления. Уже к концу прошлого столетия были составлены и опубликованы каталоги А. Перрея, А. П. Орлова, И. В. Мушкетова. К 1900 г. относится организация постоянной Центральной сейсмической комиссии, приступившей к изданию периодических бюллетеней. В 1906 г. Б. Б. Голицыным (1862—1916) организована первая сейсмическая станция в Пулкове; с 1912 г. эта станция начала систематические наблюдения. Одновременно подготавливалась, а частично и осуществлялась организация других станций. Первая мировая война приостановила эту работу.

После Октябрьской революции дальнейшее сейсмическое исследование СССР возглавил ближайший ученик и помощник академика Голицына, член-корреспондент Академии Наук СССР Павел Михайлович Никифоров (1884—1944).

С 1928 г. под его руководством начался быстрый и организованный рост сейсмологических учреждений, а с ними вместе и советской сейсмологической науки. Созданный по инициативе П. М. Никифорова в системе Академии Наук СССР Сейсмологический институт (1928) с первых же шагов повел работу в нескольких направлениях: разработка вопросов математической и физической сейсмологии, конструирование и постройка приборов, анализ методов сейсмической разведки, исследование искусственных колебательных процессов, развитие сети сейсмических станций, изучение землетрясений и составление проектов сейсмического районирования территории СССР. Последняя задача, т. е. составление схем сейсмического районирования, была решена в первом приближении к 1936 г., что оказалось возможным благодаря широкому и быстрому развитию в стране сейсмологических знаний, усовершенствованию приборов и методов обработки сейсмограмм, организации института корреспондентов-наблюдателей на местах и тщательному изучению особенностей сейсмического режима каждого района в связи с его геологическим строением.

К 1928 г. в СССР функционировало восемь сейсмических станций первого класса, цель которых — регистрация удаленных землетрясений.



456 v 457

Станции были снабжены сейсмографами системы Голицына, разработавшего принципиально новый и весьма удачный метод гальванометрической регистрации. Приборы Голицына были установлены на ряде наших и зарубежных телесейсмических станций. Для регистрации же близких и местных землетрясений эти приборы были мало пригодны отчасти вследствие своих конструктивных особенностей, отчасти вследствие того, что частотная характеристика этих приборов была подобрана Голицыным специально для регистрации удаленных землетрясений.

Для изучения близких и местных землетрясений Голицын построил тяжелые маятники с механической регистрацией, которые были установлены на ряде наших сейсмических станций. Однако и эти приборы оказались не вполне пригодными для регистрации близких и местных землетрясений вследствие малого их увеличения (около 40) и некоторых конструктивных недостатков (вибрация рычагов).

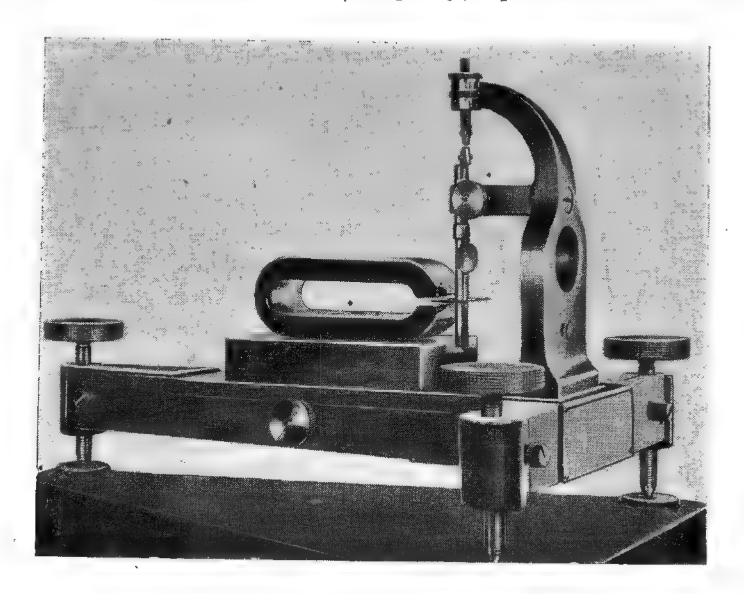
После Октябрьской революции перед отечественной сейсмологией встала задача восстановления и дальнейшего развития сейсмической службы в стране. Большинство телесейсмических станций, прекративших работу в годы первой мировой и гражданской войн, было восстановлено и пущено к 1923 г. После этого началось создание региональной сейсмической сети. Выяснилась необходимость создания аппаратуры, обладающей большим увеличением и простотой в обращении и притом свободной от искажающего влияния увеличительных рычагов. Используя оптический метод регистрации, а также маятники с короткой приведенной длиной, П. М. Никифоров создал горизонтальный сейсмограф, удовлетворяющий этим требованиям (фиг. 1).

В 1927 г. были организованы две первые региональные станции в Алма-Ата и Фрунзе; приборы П. М. Никифорова вскоре показали, что они вполне удовлетворяют предъявляемым к ним требованиям.

Период с 1928 по 1932 г. был периодом наиболее интенсивной организации региональных станций. За это время организованы станции: Андижан, Самарканд, Чимкент, Семипалатинск (в Средней Азии), Ялта, Феодосия, Севастополь, Симферополь (в Крыму), Пятигорск, Грозный, Сочи, Ереван (на Кавказе). Естественно, что организация сейсмических станций продолжалась и в дальнейшем, и к настоящему времени, в дополнение к уже упомянутым, организованы станции в Оби-Гарме и Сталинабаде (Тадж. ССР), Ленинакане (Армян. ССР); в процессе

организации находятся станции во Львове (УССР), Ключах (Камчатка), Кулябе и Мургабе (Тадж. ССР), Ашхабаде (Туркм. ССР).

Региональные сейсмические станции Средней Азии, Кавказа и Крыма доставили богатый материал, позволивший составить карту распределения эпицентров и глубин залегания очагов землетрясений на территории СССР. В некоторых случаях (в Крыму, Средней Азии и на Кавказе)



Фиг. 1.

эти данные позволили создать новые, более совершенные представления о строении земной коры и причинах землетрясений в этих районах.

Наблюдения региональных и телесейсмических станций, общее число которых доходит теперь до 30, дали основной исходный материал при построении схем сейсмического районирования. Конечно, при любом исследовании, связанном с проблемой сейсмического районирования, необходимы и неинструментальные наблюдения, дающие основной материал для суждений о бальности колебаний почвы.

Для более детального изучения природы сейсмических колебаний в последние годы была предпринята попытка проведения комплексных геолого-геофизических исследований в одном из наиболее сейсмических

районов СССР — в Таджикистане. В этих целях было необходимо в первую очередь построить приборы для записи вертикальной составляющей, т. е. вертикальные сейсмографы, пригодные для установки на региональных станциях для записи близких землетрясений (фиг. 2). Такой прибор — вертикальный сейсмограф, снабженный гальванометрической регистрацией, — в 1941 г. был разработан Д. А. Хариным. Этот прибор имеет то преимущество, что он обладает более широкополосной частотной характеристикой (его увеличение мало зависит от периодов сейсмических волн), вследствие чего его записи дают более правильное представление о движении почвы при прохождении сейсмических волн, чем сейсмограммы прежних приборов. Наблюдения, начатые с этим прибором во время Великой Отечественной войны, показали желательность подобного же улучшения в горизонтальных сейсмографах. Такие приборы были разработаны Д. П. Кирносом и построены в 1945 г. в механической мастерской Сейсмологического института АН СССР (фиг. 3). В 1946 г. эти приборы были установлены на сейсмических станциях в Сталинабаде и Оби-Гарме и успешно там работают. Большое преимущество новых систем сейсмографов заключается в их универсальности, т. е. в возможности регистрировать не только близкие, но и удаленные землетрясения.

В ряде пунктов Таджикистана ведутся в настоящее время наблюдения над режимом колебаний почвы, причем в этих исследованиях используются разнообразные приборы: новые вертикальные и горизонтальные сейсмографы системы Харина и Кирноса, наклономеры системы Бончковского, экстенсометры системы Вешнякова; проводятся по широкой программе геодезические работы (триангуляция, полигонометрия, нивелировка); ведутся гравиметрические, а также и геологические исследования. Эти комплексные работы, без сомнения, позволят детально выявить многие стороны процесса возникновения и распространения сейсмических колебаний и тем значительно уточнить и конкретизировать схематическое построение и основные положения задачи сейсмического районирования.

Параллельно с развитием сети региональных и телесейсмических станций велись исследования, непосредственно связанные с проблемой сейсмического районирования территории СССР.

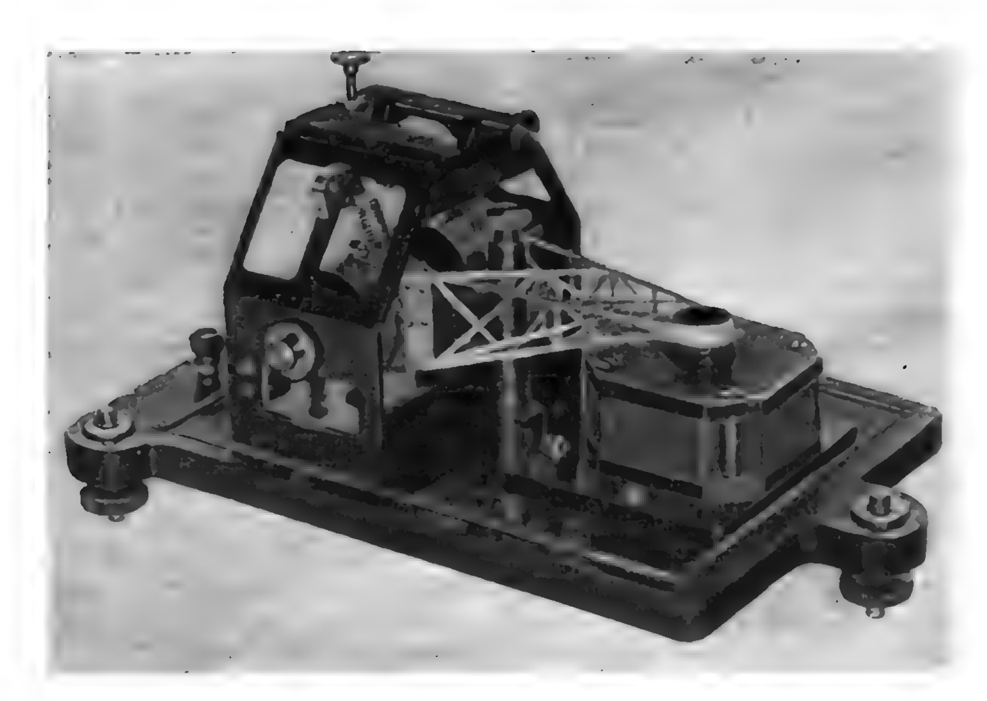
Несмотря на то, что многие районы нашей страны подвергаются частым и сильным землетрясениям, в русской литературе не было сводных

работ по вопросам сейсмичности страны в целом. В конце прошлого столетия была опубликована статья Ф. Монтессю де-Баллора о сейсмичности России, затем, вскоре после нее — статья П. Васильева для юга Европейской части СССР. В начале нашего столетия появилось много специальных исследований, давших материал для классических работ сейсмогеологического содержания (монографии И. В. Мушкетова, К. И. Богдановича, В. Н. Вебера и др.). Постепенно росло и количество каталогов землетрясений; появились отдельные каталоги для Крыма, Кавказа, Прибайкалья и т. д. Тем не менее общей сводки по землетрясениям в СССР не было.

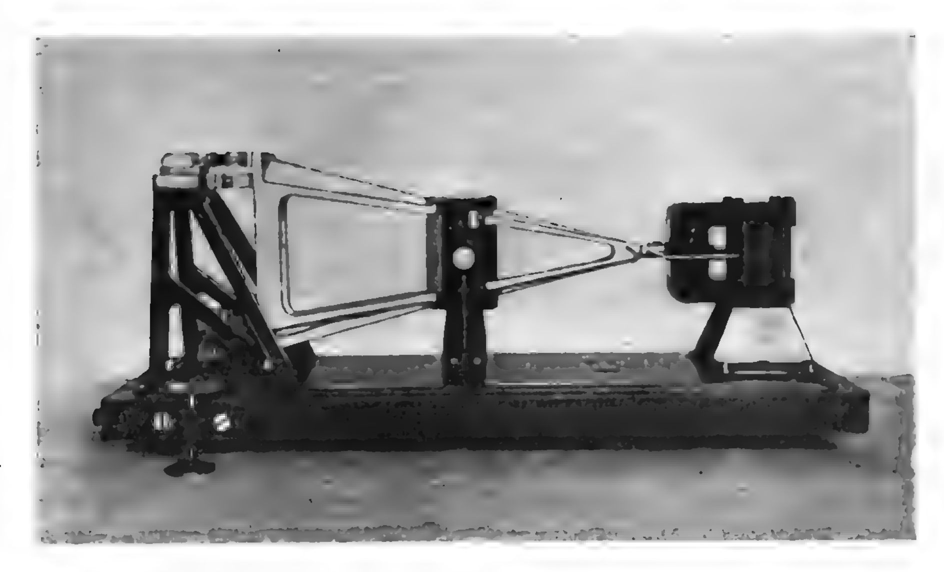
Сейсмическое районирование имеет целью создание теоретической картины распределения на данной территории будущих землетрясений; другими словами, задача заключается в решении одного из вопросов проблемы прогноза, именно вопроса о месте и силе ожидающихся землетрясений. Подобная задача может быть решена только на базе тщательного и полного анализа всего статистического материала, т. е. неинструментальных сведений о землетрясениях, материалов сейсмических станций и, наконец, данных геологии (точнее, сейсмотектоники). Таким образом, сейсмическое районирование представляет сводку всех известных к данному моменту сведений о сейсмичности данной территории и должно, следовательно, восполнить пробел в этой части в нашей специальной литературе.

Первая схема сейсмического районирования, подготовленная, как указывалось, к 1936 г., была издана в Большом советском атласе мира, а затем в ряде специальных изданий и инструкций по антисейсмическому строительству. Схема отражала общие черты сейсмического режима нашей территории, но во многом была неточна как вследствие недостатка материала (отсутствие сведений о глубине очагов и т. п.), так и вследствие непроработанности многих вопросов сейсмогеологического порядка.

За последние 10 лет удалось получить новые сведения, определеннее стали наши познания в отношении распределения эпицентров землетрясений и глубины очагов (работы А. Я. Левицкой, Н. А. Линден, Е. Ф. Саваренского, Е. А. Розовой), уточнились представления в области сейсмогеологии, получены новые сведения о важных и интересных землетрясениях и т. д. В результате появилась возможность пересмотра прежних идей и составления новой карты сейсмического районирования.



Фиг 2.



Фиг. 3.

Новая карта была составлена к 1946 г. Она в целом подобна предыдущей, но во многих своих частях, как в деталях, так и в некоторых принципиальных отношениях, отлична от нее. Карта составлена Г. П. Горшковым при ближайшем участии А. Я. Левицкой, Н. А. Линден, В. Г. Тищенко, С. В. Медведева, Е. Ф. Саваренского, А. Д. Сабининой, Е. А. Розовой, И. Е. Губина и многих других специалистов и геологов. Общее научное руководство принадлежало В. Ф. Бончковскому.

Издание новой карты сейсмического районирования территории СССР позволит точнее и строже подходить к вопросу об определении сейсмической опасности для того или иного пункта и тем самым правильнее и рациональнее распределять средства, затрачиваемые на строительные мероприятия сейсмо-профилактического порядка. Прямое назначение карты — дать возможность быстро и уверенно определять величину расчетного сейсмического балла для любого пункта СССР; другими словами, карта сейсмического районирования служит основным исходным материалом при работе по оценке и нормализации правил и мер антисейсмического строительства.

Вопросами антисейсмического строительства в России начали заниматься еще в конце XIX века в связи с рядом разрушительных землетрясений в Средней Азии. Так, после Верненского землетрясения в 1887 г., во время которого сильно пострадал город Верный (Алма-Ата), были разработаны первые в России правила строительства в сейсмических районах, предназначенные для Семиреченской области. Ряд последующих землетрясений (Красноводское в 1895 г., Шемахинское в 1902 г.) вызвал появление правил строительства для соответствующих областей, подверженных землетрясениям.

Конечно, эти правила строительства, установленные сейчас же после разрушительных землетрясений, еще не имели под собой той научной основы, которая появилась в результате систематического и углубленного изучения последствий землетрясений и анализа способов борьбы с землетрясениями в строительстве в других странах. В связи с этим была организована в 1911 г. комиссия (при обществе архитекторов-художников), в задачу которой входило систематическое изучение и разработка вопросов, связанных с антисейсмическим строительством, применение правил и норм строительства, существовавших тогда в некоторых

других странах (Италия и Япония), подготовка новых правил антисейсмического строительства и т. п. Комиссией была проведена большая работа по собиранию и объединению материалов, относящихся к землетрясениям и антисейсмическому строительству, был сделан перевод итальянской «Технической инструкции» 1909 г. и пр. Начавшаяся война в 1914—1919 гг. прервала эту работу.

Только после Октябрьской революции, в годы восстановления и развития народного хозяйства, когда во всех республиках широко развернулось новое строительство, вопросы антисейсмического нормирования возникли снова и привлекли к себе еще больше, чем раньше, внимания. Однако систематические исследования по вопросам антисейсмического строительства, в основном сосредоточившиеся в Сейсмологическом институте АН СССР и отчасти в Тбилисском институте сооружений, начались с 1928—1929 гг. Поэтому о научно-практических достижениях по вопросам антисейсмического строительства в СССР можно говорить, лишь начиная с этого момента.

Исследовательские работы в первые годы велись по двум направлениям: а) по собиранию, систематизации и анализу материалов о поведении различных сооружений и конструкций во время бывших землетрясений в СССР и за рубежом и б) по теории колебаний сооружений в применении к сейсмостойкому строительству. Результаты этих исследований были первым шагом в научной разработке принципов антисейсмического строительства в СССР. Сейсмологическим институтом был разработан проект «Инструкции по проектированию гражданских и промышленных зданий, возводимых в сейсмических районах», который после нескольких дискуссий при участии широкого круга ученых и инженеров был утвержден правительственными органами к обязательному применению всеми организациями и учреждениями СССР с 1 января 1940 г.

Наряду с этим продолжались дальнейшие научные исследования в этой области. Было проведено инструментальное изучение колебаний зданий от взрывов, уличного движения, движения поездов метрополитена, работы машин и пр., что позволило охарактеризовать динамические свойства различных зданий и их колебаний при различных возбудителях и в свою очередь могло бы послужить косвенным указанием на их поведение и во время землетрясений.

Дальнейшие обследования и анализ действия землетрясений на здания и сооружения вместе с данными инструментальных и теоретических исследований дали дополнительное обоснование теории о сейсмических воздействиях на сооружения. Было установлено, что во время сильных землетрясений в эпицентральных областях основной причиной разрушения являются вертикальные и горизонтальные составляющие импульсного характера, действие которых до тех пор недооценивалось. Их действие рассматривается как действие больших сил малой продолжительности, вызывающих концентрацию напряжений в жестких связях сооружений и нарушение этих связей. Если даже сооружение сразу и не разрушается целиком, то оно в значительной степени теряет жесткость и устойчивость, делаясь более уязвимым для последующих сейсмических колебаний (в главной фазе) более длинного периода.

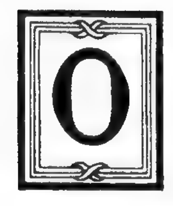
Исследования влияния вертикальных составляющих сейсмических колебаний в главной фазе также показали, что их действие может привести к быстрой потере устойчивости высоких (гибких) сооружений. В дальнейшем оказалось возможным при учете нормальной вязкости материала получить и необходимые условия устойчивости высоких сооружений во время землетрясений.

В начале Великой Отечественной войны Сейсмологическим институтом Академии Наук СССР был произведен пересмотр старой инструкции (1939 г.) и разработана совместно с Центральным научно-исследовательским институтом промышленных сооружений (ЦНИИПС) новая инструкция по антисейсмическому строительству в условиях военного времени, утвержденная Наркомстроем в 1943 г. Некоторые новые результаты исследований Сейсмологического института получили отражение в этой инструкции. В настоящее время ведется работа по уточнению инструкции и снабжению ее новой картой сейсмического районирования СССР.

Таким образом, важная проблема советской сейсмологии и проблема сейсмического районирования страны получают разрешение, а это имеет не только существенное теоретическое значение для дальнейшего развития знаний в области геофизических свойств земли, но и чисто практическое значение, связанное с планированием и распределением огромных государственных средств, расходуемых в целях предохранения жизни людей и зданий от гибели при землетрясениях.

Профессор
Н. И. Идельсон
и А. А. Немиро

КАТАЛОГ 2957 ЗВЕЗД СО СКЛОНЕНИЯМИ от -10° до $+90^\circ$



пределение координат небесных светил — одна из самых старых и важных, но вместе с тем и самых трудных проблем астрономии. Не только все наши представления о размерах, расстояниях, движениях и массах небесных тел, но и ряд задач научного и практического характера в области мореплавания, геодезии, гравиметрии и службы времени непосредственно связаны с результатами определения координат звезд, планет, Солнца и других небесных тел.

Не удивительно поэтому, что уже с давних времен подобного рода наблюдения занимают почетное место в деятельности почти всех астрономических обсерваторий мира. Благодаря блестящей постановке астрометрических работ стали всемирно известными обсерватории Пулковская, Гриничская, Вашингтонская, Капская и другие. Безукоризненно строгая трактовка астрометрических задач и удивительная точность наблюдений выделили Пулковскую обсерваторию из этого ряда знаменитых учреждений как первую среди равных.

В результате неутомимой работы астрономических обсерваторий всего мира число звездных каталогов в настоящее время очень велико (более 500), и оно продолжает непрерывно расти. Трудно представить себе ту огромную изобретательность, научную инициативу и многолетний упорный труд, которые потребовались для этого. Однако в результате

этих действительно изумляющих усилий получился хотя и обильный, но крайне разнородный материал. Прежде всего состав звездных каталогов представляет собою очень пеструю картину: число звезд в отдельных каталогах колеблется от немногих десятков до сотен тысяч, звезды включаются в каталог по самым различным признакам — яркости, спектру, расположению в той или иной области неба и т. п. Помимо того, точность отдельных каталогов очень различна, и, что самое существенное, каждому каталогу присущи свои ошибки систематического характера; как правило, они являются функциями самих координат звезд. Использование такого разнохарактерного материала представляет большие трудности.

Ввиду этого еще в середине прошлого столетия начались работы по сводке результатов многих каталогов в единую систему; подобного рода сводные каталоги получили название «нормальных или фундаментальных систем». Огромное значение этих систем заключается в том, что они доставляют нам сведения о собственных движениях звезд в пространстве, лежащие в основе наших представлений о строении и динамике мироздания. Между тем очевидно, что звездный каталог, в котором положения звезд относятся к одному и тому же моменту времени, никаких данных о движениях звезд содержать не может.

В настоящее время в качестве основного стандарта принята так называемая система FK3, т. е. «третьего улучшенного фундаментального каталога» Ауверса. По предложению Пулковской обсерватории Ауверс приступил к работе по изданию первого фундаментального звездного каталога еще в 80-х годах прошлого века; она была закончена лишь после его смерти, в 30-х годах текущего столетия. Система FK3 содержит около полутора тысяч ярких звезд, распределенных по всему небу. При составлении ее пулковские звездные каталоги сыграли очень важную роль. Существенный недостаток этой системы заключается, однако, в небольшом числе входящих в нее звезд.

Много больше звезд содержит фундаментальный каталог, составленный под руководством Беньямина и Льюиса Босса в специально созданном для выполнения этой работы институте в Америке; в нем содержится более 33 тыс. положений и собственных движений звезд.

Но, разумеется, наличие этих нормальных систем координат и движений звезд вовсе не упраздняет и не обесценивает работу обсерваторий

по составлению каталогов. Собственные движения звезд в фундаментальных каталогах по необходимости выводятся из сопоставления современных наблюдений со старыми, полученными сто или более лет назад (из использованных при составлении нормальных систем каталогов самым старым является каталог Брадлея, составленный в Гриниче в 1750—1760 гг.). Значительные, с современной точки зрения, систематические ошибки, содержащиеся в старых каталогах, передаются и выведенным при их посредстве собственным движениям звезд. Это обстоятельство самым тяжелым образом отзывается на всех работах по звездной астрономии и приводит к тому, что вычисленные на некоторую эпоху положения звезд искажены ошибками принятых для них собственных движений, и притом тем значительнее, чем дальше от средней эпохи наблюдения фундаментального каталога отстоит момент, для которого вычисляется положение звезды.

Между тем требования, предъявляемые к точности звездных положений, используемых для многих задач научного и практического характера, продолжают непрерывно расти. Первое место среди этих задач принадлежит астрономо-геодезическим работам, особенно широко поставленным в СССР.

Международный стандарт — фундаментальный каталог FK3 — не может обеспечить всех наших астрономо-геодезических работ уже в силу недостаточного числа содержащихся в нем звезд. Каталог Босса хотя и заключает в себе достаточное число звезд, но он неудобен для использования в геодезических целях ввиду необходимости производить дополнительные редукции на систему FK3. Кроме того, как уже упоминалось, положения звезд, взятые из этих фундаментальных систем, будут в некоторой степени искажены ошибками звездных движений.

Поэтому совершенно естественно возникла мысль о составлении специального звездного каталога, основанного на современных наблюдениях и содержащего достаточное для удовлетворения всех запросов геодезии число звезд. При этом было учтено то обстоятельство, что некоторое число необходимых для этой цели звезд уже имеется в каталогах, составленных из наблюдений, произведенных около 1915 г. на нескольких обсерваториях для обеспечения опорными звездами международного предприятия по составлению подробной фотографической карты неба. Наблюдения производились по специальному плану, выработанному

30*

на Международной астрономической конференции в Париже в 1909 г. Планом предусматривалось определение координат около 3000 звезд, равномерно распределенных по всему небу.

Для северного полушария неба (в зоне от +90° до —10° склонения была принята предложенная Пулковской обсерваторией программа, содержащая около 1600 звезд. В наблюдениях принимали участие обсерватории в Пулкове, Алжире, Бабельсберге, Вашингтоне, Гриниче, мысе Доброй Надежды, Николаеве, Париже и Уккле. Результаты наблюдений были впоследствии сведены по одной координате (склонение) Ленинградским астрономическим институтом, а по другой (прямое восхождение) — Пулковской обсерваторией.

Однако эти сводные каталоги не удовлетворяют всем требованиям астрономо-геодезических работ, так как в них нет некоторых очень важных в данном случае звезд промежуточной яркости, от 5-й до 6-й звездной величины.

Несмотря на то, что все эти звезды, числом около 1300, вошли в каталоги Пулковской обсерватории, составленные около 1925 г., задача обеспечения всех геодезических работ точными положениями звезд не могла считаться решенной, так как система пулковских каталогов 1925 г. отличалась от системы упомянутых выше сводных каталогов Пулковской обсерватории и Астрономического института.

На Всесоюзной астрономической конференции 1932 г. был принят выдвинутый Пулковской обсерваторией план перенаблюдения звезд пулковских каталогов 1925 г. на нескольких обсерваториях СССР. Эти звезды и получили в дальнейшем название «геодезических».

Разработанная руководителем всей работы по каталогу геодезических звезд, выдающимся советским астрометристом проф. Н. В. Циммерманом, программа наблюдений преследовала цель определения всех наблюдаемых звезд в системе упомянутых выше сводных каталогов. Иначе говоря, наблюдения были организованы таким образом, чтобы полученные в результате их обработки координаты всех определяемых 1334 звезд не отличались в систематическом отношении от положений, даваемых сводными каталогами.

В 1934 г. начались наблюдения по этому плану на пяти обсерваториях — Пулковской, Московской (ГАИШ), Казанской, Николаевской и Ташкентской; они были закончены в 1939 г. Обработка всего материала продолжалась до 1941 г. Ввиду исключительной важности этой работы

для астрономо-геодезических определений в ней приняло также участие Главное управление геодезии, съемки и картографии при Совете Министров СССР.

Всей огромной работой по созданию каталога руководил проф. Н. В. Циммерман. Его глубокое понимание задач астрометрии, неисчерпаемая энергия и богатый наблюдательский опыт проявились в полной мере в процессе выполнения этой работы и сплотили всех ее участников в дружный, согласованно работавший коллектив. Уже в блокированном Ленинграде, мужественно преодолевая все лишения, заканчивал Н. В. Циммерман работу над каталогом, и благодаря его заботам и предусмотрительности были спасены от уничтожения все ее основные результаты. После трагической смерти проф. Н. В. Циммермана каталог был подготовлен к печати в 1945 г. учениками покойного А. А. Немиро и Б. А. Орловым, внесшими в каталог необходимые и совершенно естественные добавления, не меняющие, однако, общего плана работы, намеченного ее автором.

Таким образом был составлен большой каталог 2957 звезд, объединяющий в своем составе программы сводных каталогов 1915 г. и каталогов Пулковской обсерватории 1925 г. Главное достоинство работы заключается в том, что она была выполнена коллективом пяти обсерваторий по единому и строго выдержанному плану. Благодаря этому положения всех 2957 звезд действительно относятся к вполне определенной системе. Каталог геодезических звезд содержит в себе все звезды от северного полюса до склонения —10° ярче 6-й звездной величины.

Работа по составлению каталога геодезических звезд в известном смысле подводит итоги многолетним наблюдениям ярких звезд на различных обсерваториях мира, так как в ней на основании новейших наблюдений высокой точности даются в единой и строго выдержанной системе положения всех ярких звезд северного неба. Сравнение полученного каталога с обеими фундаментальными системами (FK3 Ауверса и GC Босса) показало высокую точность нового каталога; более того, оно имеет и независимый интерес, выявляя некоторые особенности системы FK3. Каталог содержит также редукции всех звезд на систему FK3 и вспомогательные величины для приведения звезд на видимые места. Это делает весьма удобным пользование каталогом и отнесение результатов основанных на нем работ к международному стандарту — FK3.

Высокая точность и однородность системы полученного каталога позволяют расширить первоначально поставленную задачу: путем нового наблюдения в будущем на ряде обсерваторий СССР всех 2957 звезд получить высокоточную и однородную систему их собственных движений. Это не только обеспечит все геодезические работы точными положениями звезд на десятилетия вперед, но и представит глубокий и разносторонний интерес для решения многих вопросов звездной астрономии. Предварительный план этой работы встретил сочувственное отношение многих специалистов.

Таким образом, пять советских обсерваторий с успехом выполнили крупную коллективную работу, план которой был предложен и разработан в Пулкове. Ее научное и научно-практическое значение очень велико. Опыт совместной работы наших астрономических учреждений будет использован при проведении еще одной работы, предложенной и задуманной в Пулкове,— при составлении фундаментального каталога «слабых звезд». Эта работа будет одной из больших новостей в деле меридианной астрономии.

ХИМИЯ



Академик

Г. Г. Уразов

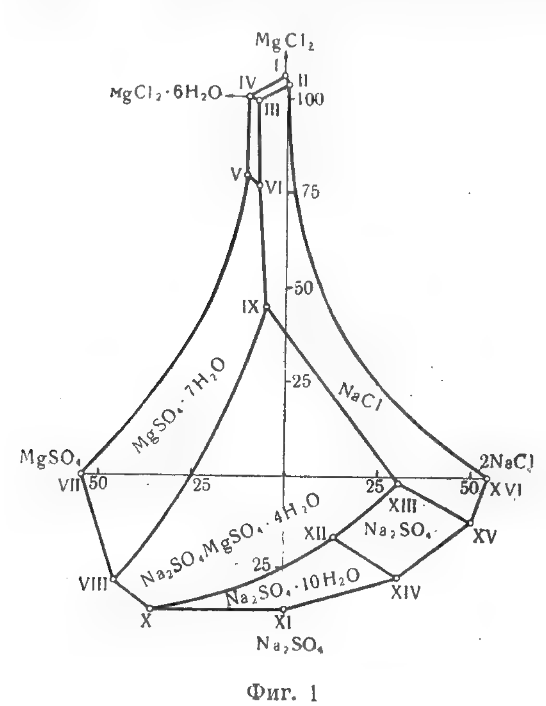
ФИЗИКО-ХИМИЧЕСКИЙ АНАЛИЗ И ЕГО ОСНОВАТЕЛЬ Н. С. КУРНАКОВ



Научные работы Николая Семеновича Курнакова в области химии охватывают обширный круг вопросов. Особенно его интересовали вопросы синтеза и строения веществ, их свойств и использования их в практических целях. В последнем направлении много внимания Н. С. Курнаков уделял использованию отечественных солевых богатств и связанному с ним изучению соляного дела. Этот вопрос его очень интересовал еще на студенческой скамье. В результате изучения соляного дела явилась его диссертация «Испарительные системы соляных варниц», представленная в 1885 г. для получения звания адъюнкта.

Этим трудом открывается длинный ряд исследований в области соляного дела —соляных водоемов и соляных отложений; ими он с увлечением занимался на протяжении всей своей жизни. Первыми объектами в этом длинном ряду исследований были одесские лиманы и крымские соляные озера. Здесь впервые Курнаковым была выявлена новая точка зрения на объекты этого рода как на природные образования, находящиеся в постоянном взаимодействии с окружающей средой и зависящие от климатических, почвенных, биологических и других факторов, обусловливающих их состояние. Состав соляных водоемов не остается постоянным, и изменения его при благоприятных условиях становятся доступными непосредственному наблюдению.

Подвергая глубокому научному анализу результаты исследований, добытых им и его сотрудниками при изучении составов соляных рассолов крымских озер, Курнаков устанавливает факт их метаморфизации и вводит новое понятие коэффициента метаморфизации природных соляных рассолов. В связи с этим большое внимание уделялось изучению явлений, происходящих в природных рассолах в связи с изменением условий их существования.



Фиг. 1

В качестве примера исследований этого рода следует остановиться на произведенном изучении четверной системы, образованной четырьмя солями $MgSO_4 + 2NaCl \rightleftarrows Na_2SO_4 + MgCl_2$ и водою. Эта система в основных своих чертах воспроизводит природные процессы, протекающие в сульфатных озерах, в частности в рапе Кара-Бугаз-гола — залива Каспийского моря, являющегося крупным сульфатным месторождением мирового значения. Полученные в результате изучения этой системы данные, представленные графически в виде диаграммы растворимости, отмечают температурно-концентрационные области существования и кристаллизации отдельных солей, образующихся в данной системе в результате протекания в ней обменной реакции

$$MgSO_4 + 2NaCl \rightleftarrows Na_2SO_4 + MgCl_2.$$

Диаграмма (фиг. 1) рисует процессы, протекающие лишь при температуре 25°, и поэтому она далеко не исчерпывает всего разнообразия изменений, происходящих в системе при других температурах. Чтобы составить представление о влиянии температуры на состояние системы, мы

должны построить соответствующие диаграммы для других температур и, пользуясь принципом непрерывности, объединить отдельные изотермические диаграммы в одно целое — политермическую диаграмму растворимости, воплощающую полноту химических превращений в рассматриваемой системе в зависимости от изменений концентрации взаимодействующих веществ и температуры.

Таким образом полученная политермическая равновесная диаграмма состояния имеет более общее значение, чем изотермические диаграммы. Она может быть применена при решении самых разнообразных вопросов использования и переработки природных сульфатных рассолов. Эта система особенно интересна в настоящее время, когда проблема естественного сульфата приобретает большую практическую важность. Пути разрешения сульфатной проблемы в значительной мере лежат через дальнейшее, более глубокое познание рассматриваемой системы. На примере этой системы может быть проиллюстрировано значение изучения равновесных диаграмм подобного рода; здесь особенно наглядно проявляется сила научного метода физико-химического анализа, разрабатывавшегося Курнаковым.

Применение физико-химического анализа к разрешению разнообразных вопросов теоретического и практического значения можно показать на многих примерах. Остановимся, однако, лишь на одном примере, связанном с проблемою советского калия.

Вся калиевая проблема и ее блестящее развитие в СССР связаны с именем Курнакова. Еще в 1916 г. им было дано заключение о калиеносности соляных отложений Соликамского района. Произведенные здесь впоследствии разведки и бурения вскрыли мощные отложения калиевых солей, ныне эксплоатируемых в широких масштабах.

В связи с открытием мощных Соликамских месторождений Курнаковым было высказано предположение об отложениях калиевых солей к югу от Северного Приуралья — до Каспийского моря. Поставленные в связи с этим экспедиционные работы подтвердили наличие калия в районе соляного озера Индер, а последующее бурение принесло многочисленные новые подтверждения нахождения здесь калийных отложений сульфатного типа (Индер, Озинки, Аще-Булак, Ак-Джар и др.). В настоящее время известно большое число точек распространения отложений калийных солей, благодаря чему можно говорить об обширной области

распространения отложений калиевых солей к востоку от Нижней Волги и на юг до Каспийского моря.

Все эти обстоятельства вызвали к жизни многочисленные исследования равновесий в системах с калиевыми солями с целью применения полученных результатов к решению вопросов, связанных с использованием калийных отложений для различного их применения, в особенности для использования их в качестве удобрений.

Собранный Курнаковым, его учениками и сотрудниками обширный экспериментальный и теоретический материал в области исследования многокомпонентных солевых систем представляет прочное, солидное научное основание для дальнейших работ в этой области. Грандиозные задачи по изучению и промышленному освоению сложных соляных месторождений, характеризующихся присутствием в них хлоридов, сульфатов, боратов и бромидов калия, магния, кальция и натрия, не могут устрашить наших галургов. Пользуясь богатым наследием Курнакова, они смело идут на овладение этой твердыни.

Помимо исследований в области соляного дела и связанных с ними соляных равновесий, Курнаков уделил много внимания исследованиям в области комплексных соединений. Полученные в этом направлении основные результаты представлены в его труде «О сложных металлических основаниях», являющемся одним из первых у нас крупных исследований в области комплексных соединений платины. Наряду с открытием ряда новых соединений платины Курнаков обнаружил чрезвычайно важную закономерность, дающую возможность при помощи реакции с тиомочевиной установить внутреннюю конфигурацию комплексных соединений двувалентной платины типа $Pt\,2AX_2$, где A — аммиак или подобное соединение (амин), а X — кислотный остаток. Закономерность Курнакова сводится к тому, что при действии тиомочевины на указанные соединения двувалентной платины циссоединения обменивают все внутрисферные заместители на тиомочевину, а в транссоединениях тиомочевина, не затрагивая аминных групп, замещает лишь кислотные остатки.

Работы Курнакова по платиновым металлам не остались без практического применения; ближайшими его сотрудниками по ленинградскому Горному институту (Н. И. Подкопаевым и Н. Н. Барабошкиным) были разработаны методы аффинажа сырой платины, осуществленные затем в промышленном масштабе на Свердловском аффинажном заводе.

Постройка этого завода позволила производить аффинаж платиновых металлов и избавила от необходимости вывозить за границу сырую платину, а оттуда импортировать очищенные металлы.

Характеризуя работы Курнакова по комплексным соединениям, академик И. И. Черняев[1] указывает, что «они сыграли колоссальную роль в развитии химии комплексных соединений как в деле опытного изучения комплексов, так и в развитии теоретических воззрений. В работах Курнакова прежде всего поражает широта постановки научных проблем и всесторонность охвата исследуемого вопроса — качество, перешедшее потом и на более поздние его работы. Он не ограничивался простым синтезом комплексных соединений, но исследовал также их физические свойства (рефракцию, кристаллическую форму, растворимость, кислотные и основные свойства), охватывая вопрос во всей его сложности. Он хорошо умел отличать свойства, имеющие важное теоретическое значение, от свойств, изучение которых имеет лишь подсобное значение при исследовании. Вопрос о показателе преломления, затронутый им в диссертации, приобрел, как известно, в связи с учением о диполях, боевое значение. Вопрос о кислотных и основных свойствах комплексных соединений в зависимости от природы внутрисферных заменителей имеет также принципиальное значение, и решение его принадлежит будущему».

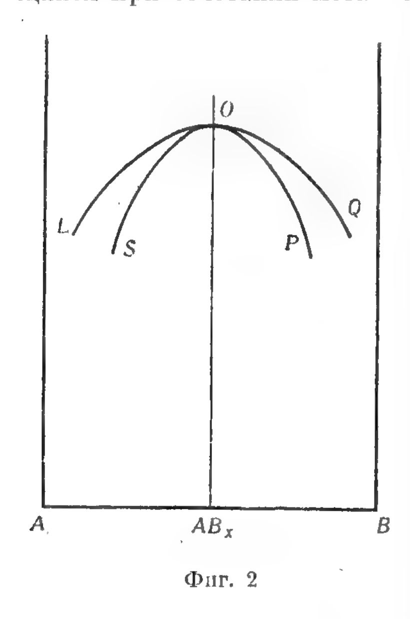
Дальнейшие работы Курнакова связаны с изучением металлических сплавов, с развитием учения о химической диаграмме «состав — свойство», с физико-химическими исследованиями природных образований и т. д.

Работами Курнакова в области изучения металлических сплавов открывается новая блестящая страница в истории развития металлографии не только у нас, но и в других странах. Построенные на широких научных обобщениях, опирающихся на периодическую систему элементов Менделеева и на бурно тогда развивавшееся учение о растворах и гетерогенных равновесиях, эти работы Курнакова не только дали ряд очень ценных и важных в научном и практическом отношении закономерностей, но и привели его к глубоким обобщениям общетеоретического характера: определению понятия химического индивидуума, развитию учения о химической диаграмме «состав — свойство» и созданию на этой почве особого

[1] И. И. Черняев. Изв. Сектора физ.-хим. анализа, 1941, т. 14, стр. 141.

отдела общей химии — физико-химического анализа, являющегося в настоящее время мощным орудием в разрешении самых разнообразных вопросов теоретического и практического характера.

В раннем периоде развития металлографии известные трудности представляло разрешение вопроса о количественном составе веществ, образующихся при сочетании металлов. Рассмотрению этого вопроса посвящена работа Курнакова «О нахождении состава определенных соединений в сплавах по методу плавкости»,[1] где впервые затрагиваются вопросы большой принципиальной важности, приведшие его впоследствии к определению понятия химического индивидуума. В статье рассматривается случай, когда соединение AB_x способно растворять в твердом состоянии переменные количества своих компонентов A и B. «В этом случае, — говорит Курнаков, — отношение между последними (компонентами. — *Г. У.*) в твердой фазе или твердом растворе будет непрерывно изменяться в зависимости от концентрации соответствующего жидкого сплава. При таких условиях, конечно, нечего говорить о неизменном значении x в формуле AB_x, которое удовлетворяло бы законам постоянства состава и кратных пропорций.



Фиг. 2

В том случае, когда тело AB_x обладает максимумом плавкости, изменения в составе жидкой и твердой фаз могут быть изображены посредством двух кривых линий LOQ и SOP (фиг. 2), касающихся друг друга в максимальной точке O, в которой состав обеих фаз одинаков.

Теперь спрашивается, будет ли удовлетворять величина коэффициента x в максимальной точке O закону кратных пропорций, или, другими словами, можно ли считать тело AB_x за определенное химическое соединение, способное давать твердые растворы со своими составными частями?

[1] Записки Русск. техн. об-ва, 1901, т. 35, № 1/7.

Решение этого вопроса может быть произведено только экспериментальным путем. Так как при условиях, принятых выше, величина x для твердого тела изменяется непрерывно, то возможно допустить случаи, когда x будет выражено числом, которое не отвечает простому кратному отношению. До сих пор еще не установлено с достаточной точностью существование подобных отношений, потому что имеющиеся в литературе примеры требуют более детальной экспериментальной проверки».

Для осуществления эспериментальной проверки указанного вопроса, а также в целях более детального и глубокого изучения природы металлических сплавов начинаются многочисленные работы Курнакова по разработке новых методов исследования и усовершенствования существующих.

В первую очередь был применен метод электропроводности. Полученные экспериментальные результаты дали возможность вывести типические диаграммы электропроводности простейших бинарных систем, в состав которых входят металлы, способные давать твердые растворы.

В тесном соотношении с электропроводностью находится другое очень важное и характерное свойство металлических сплавов — их твердость.

С течением времени в круг изучения стали вводиться и другие свойства сплавов: давление истечения, коэффициент линейного расширения, температурный коэффициент электропроводности, тепловое расширение, механические испытания, электродвижущие силы, магнитные свойства и пр.

Наряду с разработкой методов исследования значительное внимание Курнаков уделяет в первых же своих работах по металлическим сплавам также и выяснению закономерностей при взаимных соединениях металлов. Этот последний вопрос представляет не только большой теоретический интерес, но и не меньшую практическую важность. Курнаков приходит к заключению, что область взаимных металлических сочетаний должна заключать в себе особые, свойственные для них типы, характеризующиеся своей стойкостью и способностью повторяться в различных рядах веществ. Эти выводы блестяще оправдываются изучением диаграмм плавкости целого ряда бинарных металлических систем, в особенности тех, которые представляют сочетания щелочных и щелочноземельных

металлов, являющихся главными носителями основных или металлических свойств.

Совокупность полученных в этих исследованиях результатов приводит Курнакова к положению, что соединения данного металлического элемента с другими металлами обнаруживают несомненную аналогию с типами соответствующих водородистых производных.

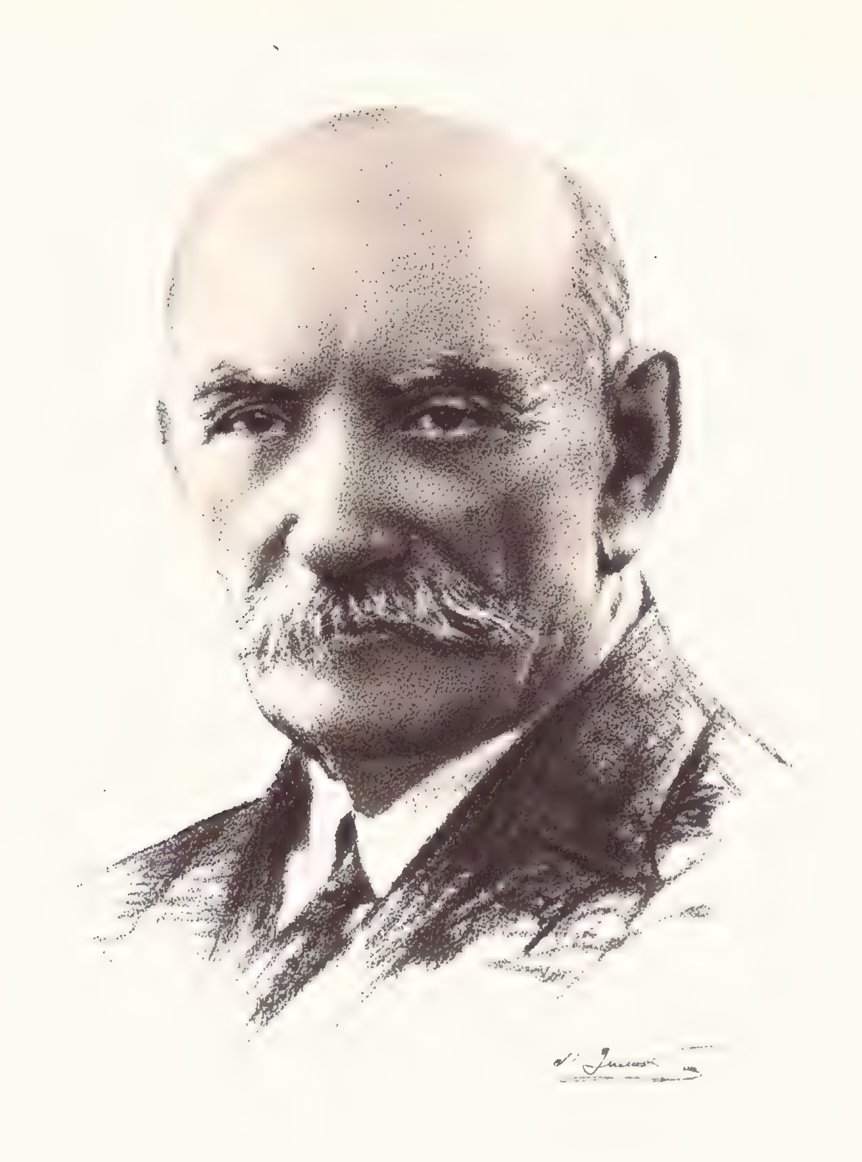
Все работы Курнакова по металлическим сплавам характеризуются одной особенностью: все они — пример сочетания глубокой теории с насущнейшими вопросами практики. Классификация металлидов на соединения бертоллетовского и дальтоновского типов, установление сингулярных элементов химической диаграммы и нахождение зависимости между свойствами и составом равновесных систем одинаково важны как для теории металлических сплавов, так и для практического применения их в различных областях техники.

Установление закономерностей между физико-химическими свойствами металлических сплавов и их составами — одна из задач познания природы металлических сплавов, задач, имеющих важное практическое значение. Это создает для практика-металловеда возможность целеустремленного, рационального подбора металлических сплавов с нужными эксплоатационными качествами. На эту сторону исследований металлических сплавов Курнаков обращал очень большое внимание. В его исследованиях с очевидностью показана связь, существующая между различными физическими и механическими свойствами металлических сплавов, особенно по изучению давления истечения и твердости пластических тел. Вот что он сам говорит об этом:

«Изучение электропроводности металлических сплавов проливает новый свет на соотношения между химическим составом и изменениями свойств равновесных систем вообще...

Действительно, электропроводность находится в тесной связи с теплопроводностью, термоэлектрической силой, а также с целой группою свойств, зависящих от внутреннего сцепления твердого вещества. Сюда следует отнести твердость, предел упругости, различные модули упругих деформаций, коэффициенты сжимаемости, временного сопротивления разрыву и другие, так называемые механические свойства...

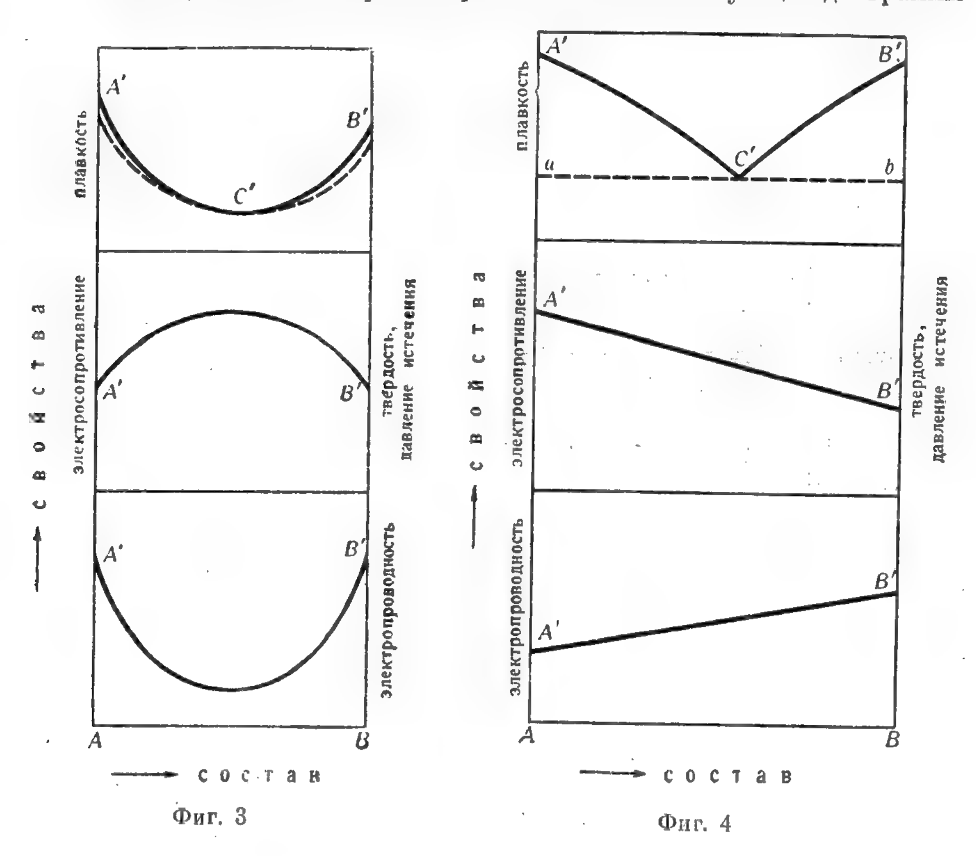
Во главе механических свойств можно поставить твердость, для которой соответствие с электропроводностью металлических систем



480 и 481

является установленным рядом экспериментальных работ последнего времени.

Если известны изменения электропроводности в зависимости от состава системы, то можно заранее предвидеть соответствующие диаграммы



Фиг. 3

Фиг. 4

твердости, а вместе с ними и другие «механические характеристики» твердого тела».

Среди механических свойств, имеющих вообще большое значение в оценке пригодности металлических сплавов для различных практических целей, особенное положение занимает внутреннее трение. Определение последнего методом давлений истечения, разработанным Н. С. Курнаковым и С. Ф. Жемчужным, и установление соотношений между давле-

нием истечения и твердостью пластических тел являются научным основанием для дальнейших исследований в этой очень важной в практическом отношении области.

Установление закономерностей между электросопротивлением металлических сплавов и их составом, а также выяснение факта влияния образования твердых растворов на понижение электропроводности и ее температурного коэффициента сыграли огромную роль в дальнейшей судьбе развития техники получения реостатных сплавов. Нахождение новых сплавов, обладающих высоким электросопротивлением и ничтожным, почти нулевым, температурным коэффициентом, становится с этих пор предметом не грубого эмпиризма, а научного исследования.

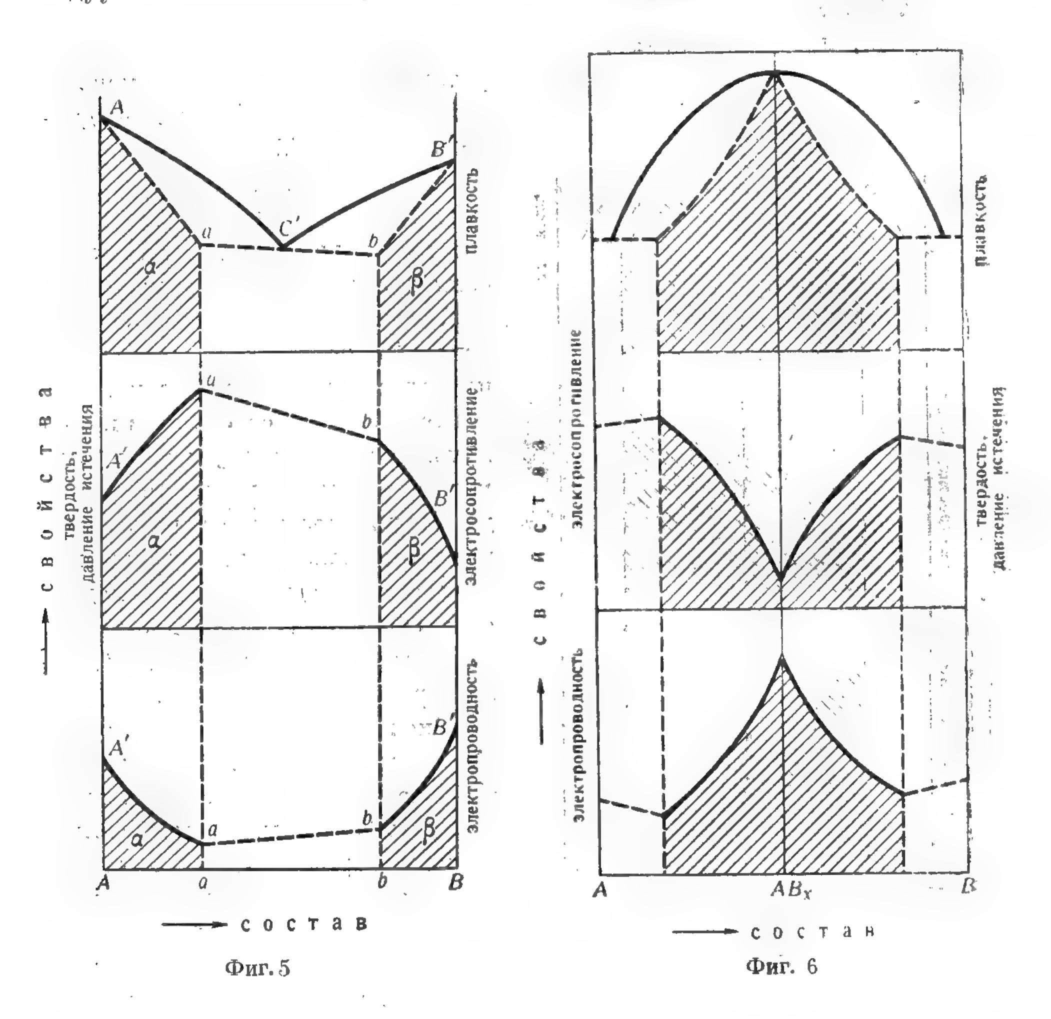
Главное внимание при изучении металлических сплавов Курнаков сосредоточивает на таких сочетаниях металлов, где наблюдается образование твердых растворов, представляющих наибольший интерес и в практическом отношении. Особенно существенны при этом явления, сопровождающие распад твердых растворов. И здесь первые систематические и детальные исследования принадлежат Курнакову и его сотрудникам.

Особое внимание Курнакова, как уже отмечалось, привлекают металлические сочетания, образующие твердые растворы и фазы переменного состава. Исследования в этом направлении привели его к установлению понятия химической диаграммы «состав — свойство», ее сингулярных элементов и к определению понятия химического индивидуума.

Обширный экспериментальный материал по изучению физических свойств твердых металлических растворов показал, что образование их сопровождается увеличением твердости, давления истечения, электросопротивления и, наоборот, понижением электропроводности. Вид кривых изменения этих физических свойств металлических сплавов в зависимости от их состава приведен на диаграммах (фиг. 3, 4, 5), где по оси абсцисс отложен состав сплавов, а по оси ординат — свойства.

Открытые соотношения между химическим составом и указанными физическими свойствами металлических сплавов дали в руки исследователя мощное орудие для определения таких тонких различий в состоянии изучаемых тел, которые были совершенно недоступны для обычно тогда применявшихся приемов химического исследования. Особенно продуктивным оказалось применение указанных соотношений для разре-

шения поставленного Курнаковым вопроса о природе химического индивидуума в связи с обсуждением случая, графически представленного



Фиг. 5

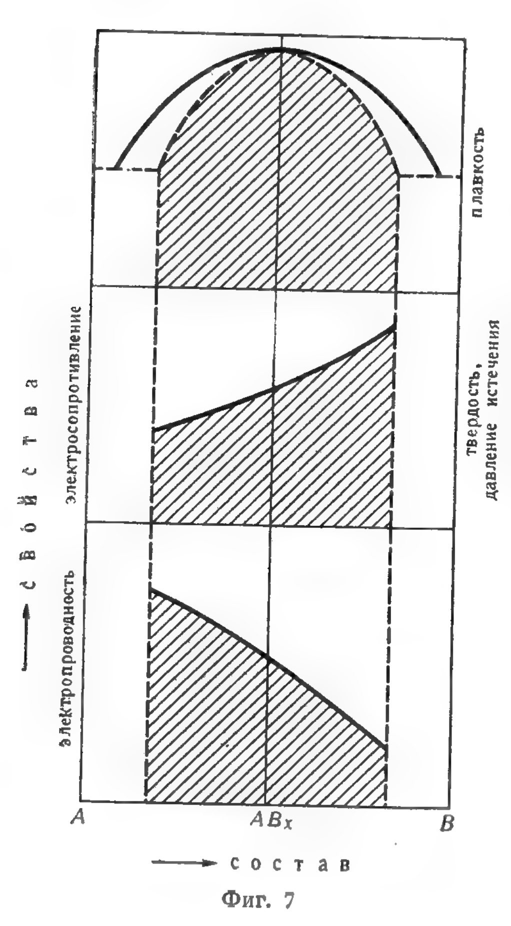
Фиг. 6

на фиг. 2. Здесь для коэффициента x возможны два случая: один — когда x выражается числом, отвечающим законам простых и кратных пропорций, и другой — когда x не выражается таким числом. Для обоих случаев, как показали экспериментальные наблюдения, картина изменения

31*

физических свойств соответствующих однородных твердых тел оказывается совершенно различной, именно такой, какой она представлена на фиг. 6 и 7.

Случай, представленный на фиг. 6, впервые экспериментально был установлен на примерах моноаргентида магния MgAg и монокадмида магния MgCd.



Фиг. 7

Случай, представленный на фиг. 7, впервые был констатирован при изучении систем таллий — свинец и таллий — висмут.

Рассмотрение общего случая (фиг. 2) результатов исследований аргентида и кадмида магния, сплавов таллия со свинцом и висмутом, а также меркуридов и кадмидов щелочных металлов приводят Курнакова к заключению, что характеристика определенного химического соединения не находится в обязательной связи с постоянством концентраций твердой фазы в равновесной системе.

Число реальных примеров общего случая (на фиг. 2) среди металлических сочетаний оказывается очень большим. В связи с этим Курнаков указывает, что «этого рода фазы являются представителями соединений переменного состава, на существование которых указал Бертолле в своем споре с Пру в начале прошлого столетия. В истории химии обыкновенно принято считать, что названный спор закончился победой Пру, установившего в науке закон постоянства состава. Несомненно, эта победа была лишь временной.

Теперь, столетие спустя, мы приступаем к разрешению тех же вопросов, которые волновали современников Бертолле и Пру, вооруженные новыми методами экспериментального исследования. Обладание методами физико-химического анализа доставляет нам возможность заняться систематическим изучением именно тех областей, которые были уже указаны Бертолле, но оставались в течение долгого времени совершенно недоступными для обычных приемов химического наблюдения».

В результате многочисленных позднейших изучений наряду с соединениями бертоллетовского типа — «бертоллидов», характеризующимися иррациональными максимумами, Курнаковым с сотрудниками были установлены и определенные химические соединения, обладающие на диаграммах плавкости рациональными температурными максимумами, у которых состав твердой фазы подвергается колебанию в более или менее широких пределах. Такого рода фазы получили название «дальтонидов». Приведенные примеры наглядно показывают всю важность и значение изучения физических свойств систем в зависимости от их состава и построения соответствующих диаграмм «состав — свойство». Последние с особенной ясностью вскрывают характер происходящих в системе химических изменений — образования новых веществ — и устанавливают природу последних.

Значение диаграммы «состав — свойство» в практическом применении можно сравнить с топографическими или геологическими картами. Подобно тому как последние являются необходимой руководящей основой при осуществлении ряда важных военных, народнохозяйственных и исследовательских мероприятий, диаграммы «состав — свойство», имея с картами много общего, дают ключ для правильного понимания и направления изучаемого процесса. Особенно ясным это становится в отношении одного из видов диаграмм «состав — свойство», именно диаграмм состояний.

В результате многолетних и обширных работ, проведенных совместно с учениками и сотрудниками в области исследования диаграмм «состав — свойство» систем, образованных самыми разнообразными веществами (металлами, солями и органическими соединениями), Курнаков приходит к убеждению о необходимости создания в комплексе современных физико-химических знаний новой области — физико-химического анализа.[1] Под

[1] Н. С. Курнаков. Введение в физико-химический анализ. Изд. IV, М., 1940.

физико-химическим анализом он подразумевал новую область химического исследования, новый, самостоятельный раздел общей химии, «который имеет целью определение соотношений между составом x и свойствами ε равновесных систем, результатом чего является графическое построение диаграммы «состав—свойство» (x, ε)». Физико-химический анализ вырос, как он сам говорил, из запросов практической металлографии.

Одна из основных черт физико-химического анализа заключается в том, что он представляет собой определенный научный метод рассмотрения химических явлений, представляет точку зрения, с которой исследователь подходит к изучению химических превращений и веществ, участвующих в них. Этот научный метод, эта точка зрения состоят прежде всего в том, что они указывают на невозможность получения правильного представления о характере и свойстве веществ, которые синтезируют химики и с которыми они имеют дело, на основе одного лишь изучения вещей в самих себе. Правильное представление о природе и свойствах веществ мы получаем не столько путем изучения самого вещества как такового, сколько путем исследования его в процессе взаимодействия с другими веществами в условиях непрерывного изменения факторов, влияющих на ход процесса.

Из возможных путей осуществления такого приема в практике химических исследований путь изучения равновесных систем, образованных совокупностью тел, способных к химическому взаимодействию, должен быть принят во внимание прежде всего. По словам Курнакова,[1] «их изучение (равновесных систем.—*Г. У.*) является одной из важнейших задач общей химии. Такие системы находятся под влиянием различных факторов; из них температуру и давление можно назвать внешними, а концентрации реагирующих веществ — внутренними факторами, обусловливающими направление превращения в ту или другую сторону».

Так как исследование превращений, протекающих в системах, мы можем осуществлять лишь путем изучения свойств систем, то «общий прием физико-химического анализа состоит в количественном изучении свойств равновесных систем, образованных двумя и более компонентами, в зависимости от их состава».[2]

---

[1] Н. С. Курнаков. Введение в физико-химический анализ. М., 1940, стр. 77.
[2] Там же, стр. 12.

Числовые результаты экспериментальных изучений свойств равновесных систем представляются обычно построением соответствующих диаграмм «состав — свойство», которые и являются тем средством, при помощи которого мы познаем вещество, условия его образования и взаимодействия.

Представляя собою определенный геометрический комплекс точек, линий, поверхностей и объемов, химическая диаграмма «состав —свойство» своими геометрическими признаками позволяет не только наглядно показать появляющиеся в изучаемой системе вещества, характеризовать их физико-химические свойства, определять границы их устойчивости и законы их совместного сосуществования, но и заключать о характере химических превращений, протекающих в системе. Вот почему построение химических диаграмм «состав—свойство» приобретает большое значение в самых разнообразных областях теоретического и практического знания, где приходится рассматривать вопросы свойств и химического взаимодействия в системах, образованных различного рода веществами. Совершенно прав Н. И. Степанов, утверждая, что химическая «диаграмма представляет собой, таким образом, наше орудие производства, наше оружие, с которым мы идем в бой за познание вещества, а следовательно, за овладение веществом».[1]

Широкое применение химической диаграммы «дает право говорить об общем геометрическом методе изучения превращений веществ, так как здесь, — говорит Н. С. Курнаков, — мы имеем пример корреляции, соответствия между химическими процессами, протекающими в системе, и геометрическими свойствами диаграммы «состав — свойство». По геометрическим изменениям в строении комплекса мы можем делать заключения о соответствующих химических превращениях, о появлении новых соединений и т. п. В этом проявляются принципы качественной геометрии или топологии, дающие надежное основание для построения химических диаграмм и облегчающих практическое применение ее при изучении природных и технологических процессов».[2]

Физико-химический анализ в настоящее время находится в одной из стадий своего развития. Основной объект его изучения — равновесные

[1] Н. И. Степанов. Метрика равновесной химической диаграммы. Усп. химии, 1936, т. 5, вып. 7—8, стр. 18.

[2] Н. С. Курнаков. Введение в физико-химический анализ. М., 1940, стр. 83.

системы, образованные одним, двумя и бо́льшим числом компонентов, путем исследования соответствующих диаграмм «состав — свойство».

Изучение последней как самостоятельного геометрического образа — одна из главных задач физико-химического анализа, осуществляемых в двух направлениях. Во-первых, качественное изучение общих геометрических свойств диаграммы, неизменных при ее преобразованиях, т. е. топологии химической диаграммы, и, во-вторых, разрешение вопроса построения теоретической диаграммы во всем ее многообразии, исходя из основных законов, управляющих химическими превращениями. Последние исследования стремятся установить количественную связь между элементами диаграммы и величинами, характеризующими химический процесс, протекающий в системе.

Применение закона химического действия масс к разрешению вопроса о прочности, степени диссоциации химического соединения, возникающего в растворах, привело Н. И. Степанова к заключению, что этот закон развертывает перед нами полную картину химического процесса, протекающего в системе.

Установление связи между формой изотермы реакции и различными численными значениями для константы реакции рисует эволюцию диаграммы при изменении прочности возникающего соединения и показывает, что в случае полного отсутствия диссоциации все непрерывные кривые (гиперболы) претерпевают вырождение и распадаются на пучки прямых. Эти пучки в пересечении дают химическую сингулярную точку, получившую название «точки Курнакова» в честь Н. С. Курнакова, являющегося родоначальником понятия сингулярных элементов химических диаграмм. Здесь мы встречаем яркий пример, подтверждающий одно из основных положений топологии, а именно пример соответствия между химическим процессом, протекающим в системе, и геометрическими свойствами диаграммы. Благодаря этому принципы качественной геометрии или топологии дают надежное основание для построения химических диаграмм и облегчают практическое применение их при изучении природных и технологических процессов.

Ввиду того, что последние характеризуются обычно участием в них большого количества веществ, часто сильно разнящихся по своим физико-химическим свойствам, перед теорией физико-химического анализа в самом начале его возникновения встали вопросы разработки графиче-

ских и экспериментальных методов исследования сложных, многокомпонентных систем. Успешному разрешению первой задачи много способствовал В. П. Радищев, разработавший методы изображения многокомпонентных систем с помощью проекции фигур в пространстве многих измерений. Приложение этого метода к реальным объектам изучения, как показали многочисленные исследования А. Г. Бергмана и его сотрудников, оказалось очень целесообразным и ценным. Это в значительной степени способствовало изучению топологии сложных химических диаграмм и установлению связи между геометрическими, термодинамическими и химическими свойствами многокомпонентных систем. В этом направлении исследованы основные типы пятикомпонентных систем, систем из восьми и более (до десяти) солей. Эти работы ИОНХ далеко опередили аналогичные работы за рубежом; тем не менее оба указанные направления в области физико-химического анализа представляют еще обширное поле для дальнейших изысканий.

Как следует из сказанного, физико-химический анализ является дальнейшим развитием современного периода учения о химических равновесиях, начатого В. Гиббсом и развивавшегося ван-дер-Ваальсом, Розебомом, Ле-Шателье, Вант-Гоффом, Тамманом и др.

Широко применяемый в физико-химическом анализе геометрический метод изучения превращения веществ в многокомпонентных системах значительно способствовал разрешению многих сложных теоретических и практических вопросов. Рассмотрение в системах превращений, протекающих под влиянием как внешних факторов равновесия (температуры, давления и др.), так и внутренних (компонентов системы), придает изучению полноту, обеспечивающую всестороннее освещение исследуемого объекта. Эта черта придает физико-химическому анализу как научному методу исследования химических явлений большую силу и могущество при решении самых разнообразных теоретических и практических задач. Если принять, кроме того, во внимание, что в ряде исследований химических систем приходилось вводить как одну из характеристик вещества еще и время, то это делает физико-химический анализ еще более универсальным.

Крупная роль физико-химического анализа как определенного объекта химического знания и как научного метода познания превращений и свойств веществ проявляется не только в аспекте исторического развития ос-

новных идей химии, но и, пожалуй, с логической точки зрения. Основной геометрический образ, применяемый физико-химическим анализом в качестве главного орудия познания химических явлений, — химическая диаграмма «состав — свойство» включает в себе основные понятия, категории, позволяющие любые явления рассматривать всесторонне, во всей их полноте, со всеми их связями. Эти категории суть: вещество, качество и отношение. Им в химической диаграмме «состав — свойство» отвечают: компоненты, свойства и изменение последних при непрерывных и последовательных изменениях отношений компонентов.

Вот почему дальнейшее развитие физико-химического анализа как определенного объекта химического значения и как научного метода исследования химических явлений представляет важную и плодотворную задачу.

Академик

Н. Н. Семенов

ХИМИЧЕСКАЯ КИНЕТИКА И ТЕОРИЯ ГОРЕНИЯ



емногим более полустолетия назад работами Вант-Гоффа и Аррениуса были заложены основы учения о скоростях химических реакций. Была дана первая классификация химических реакций простых типов и выяснена природа зависимости скорости химических реакций от температуры.

Однако уже сам Вант-Гофф понимал, что простые закономерности присущи лишь очень ограниченному числу реакций и что реальный химический процесс сильнейшим образом осложнен различными влияниями («возмущающими действиями»), сводящими на-нет простые представления о механизме химического процесса. К сожалению, в течение последующих двух-трех десятилетий ученые, работавшие над вопросами химической кинетики, искусственно ограничивали поле своих исследований поисками и изучением реакций, подчиняющихся простым закономерностям. И только в последние десятилетия при активном участии советских ученых в центре внимания химической кинетики встал реальный химический процесс со всеми сопровождающими и осложняющими его химическими и физическими явлениями.

РОЛЬ РУССКОЙ НАУКИ В РАЗВИТИИ ХИМИЧЕСКОЙ КИНЕТИКИ

Ученые нашей страны с первых же дней развития науки о скоростях и механизме химических реакций идут в первых рядах исследователей этой области.

За несколько лет до работ Вант-Гоффа русский химик Н. А. Меншуткин начал изучать скорости химических реакций. Исследования Меншуткина касались таких важных вопросов, как установление связи между строением веществ и скоростью их превращения, действие среды на химическое превращение, начальное ускорение некоторых реакций и пр. Результатами опытов Меншуткина широко пользовался Вант-Гофф для обоснования своих представлений о механизме химических превращений. Так называемые «меншуткинские реакции» стоят до сих пор в центре внимания химической кинетики. Коновалов почти одновременно с Оствальдом и независимо от него установил закон автокатализа. Бах явился творцом перекисной теории окислительных процессов. Эта теория сыграла основную роль в развитии современных представлений о механизме окислительных реакций и действии промежуточных продуктов. Шилов в начале нашего столетия выполнил ряд выдающихся работ по изучению процессов сопряженного окисления, в которых подчеркнул определяющую роль промежуточных продуктов, возникающих в ходе химического превращения.

Начиная с 20-х годов нашего столетия, началось интенсивное развитие химической кинетики советскими учеными. В области химической кинетики у нас работают и физики (Теренин, Семенов, Кондратьев и др.), и физико-химики (Фрумкин, Рогинский, Сыркин, Бродский, Кобозев, Нейман, Ковальский и др.), и химики-органики (Зелинский, Баландин, Медведев и др.), и биохимики (Парнас, Опарин, Энгельгардт и др.). Особенное развитие приобрело изучение процессов горения и взрывов на основе химической кинетики (Зельдович, Харитон, Беляев, Соколик, Франк-Каменецкий, Щелкин и др.).

Широки и разнообразны направления, в которых работают советские ученые различных школ в области химической кинетики. В дальнейшем мы остановимся только на одном из этих направлений — на работах школы Института химической физики.

ЦЕПНЫЕ РЕАКЦИИ

Основные успехи нашей школы связаны с развитием теории сложных химических процессов, теории цепных разветвленных химических реакций и с созданием теории процессов горения и взрывов на основе химической кинетики.

Значительное внимание было уделено изучению элементарных химических процессов, составляющих сложную химическую реакцию, и установлению активных промежуточных продуктов, ответственных за ход реакции.

Ведущую роль здесь сыграла цепная теория химических реакций. Первоначальная идея цепной реакции появилась в науке в 1913 г. Однако вплоть до 1927 г. эти реакции рассматривались как редкое исключение.

Картина существенно изменилась после работ советских ученых, начатых в 1926 г., в результате которых были экспериментально обнаружены так называемые разветвленные цепные реакции, открыто и объяснено фундаментальное явление предела.

В период с 1926 по 1932 г. советские ученые (Харитон, Семенов, Рябинин, Трифонов, Загулин, Ковальский, Чирков, Садовников, Нейман и др.) развивали теорию разветвленных цепных реакций, множили экспериментальный материал по изучению и обнаружению новых случаев такого рода.

Начиная с 1928 г., аналогичные работы были начаты также в Англии школой Хиншельвуда (Хиншельвуд, Диксон, Лудлам, Мельвилль, Норриш и др.).

В настоящее время можно утверждать, что огромное большинство окислительных реакций, по крайней мере в газовой фазе, протекает по механизму разветвленных цепных реакций. Такой механизм имеет место и при полимеризации многих органических веществ, как это особенно отчетливо показано Медведевым и его сотрудниками в Институте имени Карпова.

Очень вероятно, что цепные реакции несравненно более распространены в органической химии, чем обычно считают. Пора было бы поставить гораздо шире вопрос о роли радикалов и цепных реакций в биологических процессах.

Представление о разветвленной цепи позволило объяснить многие загадочные химические явления, совершенно непонятные с точки зрения классической кинетики Вант-Гоффа — Аррениуса. И более того, обнаружение этих явлений может служить признаком цепного разветвленного механизма изучаемой реакции.

ПРЕДЕЛЫ

Удивительное явление предела протекания реакции заключается в том, что реакция идет по принципу «все или ничего». В некоторых условиях реакция совершенно не идет. Но достаточно ничтожного изменения температуры или давления, чтобы химический процесс начал развиваться с громадной, неизмеримой скоростью. Это явление, открытое в Ленинграде на примере окисления фосфора, было наблюдено для многих окислительных реакций, идущих с возникновением пламени, причем был отчетливо показан нетепловой характер явления.

Работами английских (Норриш и Спенс) и советских (Садовников, Шантарович) ученых показано, что пределы бывают и в медленных реакциях. В особенно демонстративной форме это было обнаружено Шантаровичем для окисления мышьяковистого водорода. Но в действительности распространенность и значение этого явления гораздо шире. Оно отнюдь не ограничено областью окислительных реакций. Так, совершенно аналогичные явления были обнаружены Апиным при разложении хлористого азота. Своего апогея понятие предела достигло в явлениях ядерных превращений, где аналогичные явления, как недавно было показано, лежат в основе получения атомной энергии.

Исчерпывающее объяснение явлений предела с точки зрения цепной теории — одна из выдающихся заслуг советской школы в области химической кинетики. Это объяснение многократно уже мною излагалось, и нет нужды повторять его здесь.

АВТОКАТАЛИЗ И ПЕРИОД ИНДУКЦИИ

В связи с теорией разветвляющихся цепей получило свое окончательное объяснение прогрессивное самоускорение газовых реакций. Семенов, Чирков, Ковальский, Садовников нашли, что целый ряд окислительных реакций ускоряется со временем по одному и тому же экспоненциальному закону (аналогично обычному закону автокатализа). Вначале, пока активные продукты не успели накопиться, скорость реакции настолько мала, что ее невозможно заметить — это так называемый период индукции. При поверхностном наблюдении кажется, что в периоде индукции вообще ничего не происходит, что протекание реакции зависит не только

от физических условий и химического состава реагирующей смеси, но и непосредственно от времени. Но более глубокое исследование показывает, что в периоде индукции происходит накопление в реагирующей системе активных промежуточных продуктов. Наиболее детально и тщательно накопление их было установлено в обширных исследованиях Неймана и его сотрудников (1936—1939) над окислением углеводородов и Эмануэля (1939—1942) над окислением сероводорода. В первом случае роль активных промежуточных продуктов играют кислородсодержащие органические соединения — органические перекиси и альдегиды, во втором — моноокись серы.

Явления автокатализа и периода индукции также отнюдь не ограничиваются областью окислительных реакций. В работах Медведева и его школы эти явления были обнаружены и исследованы для ряда процессов полимеризации.

Эти работы сделали очень вероятным предположение, что многие реакции полимеризации идут по цепному механизму с разветвляющимися цепями. Недавно Франк-Каменецкому (1943) удалось обнаружить период индукции для такой классической органической реакции, какой является образование ароматических углеводородов из ацетилена. Период индукции чрезвычайно сильно возрастает при добавлении малых количеств окиси азота; отсюда приходится сделать неожиданное заключение, что и эта реакция идет через посредство свободных радикалов.

ПРОМЕЖУТОЧНЫЕ ПРОДУКТЫ

Общая теория рассмотренных явлений могла быть развита вне связи ее с конкретными носителями цепи, без установления детального химического механизма соответствующих реакций. Однако задача построения конкретного механизма реакций основных типов является чуть ли не самой главной задачей химической кинетики. В неразветвленных цепных реакциях, подобных фотохимической реакции хлора с водородом или термическому разложению органических соединений, иностранным ученым уже довольно давно удалось с помощью остроумных приемов доказать образование свободных атомов водорода (в первом случае) и метильных радикалов (во втором). Однако эти работы были недостаточно прецизионны; они носили только качественный характер.

В реакциях с прямыми неразветвленными цепями нельзя было теоретически ожидать высокой концентрации активных центров. По-иному обстоит дело в реакциях с разветвляющимися цепями. Недавно Семенов указал на необходимость возникновения больших количеств активных полупродуктов типа свободных радикалов, которые должны образовываться в ходе процесса с разветвленными цепями в количествах, значительно превышающих равновесные термодинамические значения, отвечающие условиям опыта. Больших успехов в деле установления химического механизма конкретных разветвленных цепных реакций достиг Кондратьев.

Кондратьев и его сотрудники (1936—1944) с помощью усовершенствованного спектрального метода установили, что в разреженном пламени водорода при температуре около 1300°С концентрация свободного гидроксила может достигать 0.5—1% от исходного водорода. Это количество в сотни раз превышает термодинамически равновесную величину, что служит одним из сильнейших доказательств цепной теории. Изучив количественно зависимость концентрации ОН от скорости реакции, Кондратьев (1944) увязал эту зависимость с цепной схемой реакции. Еще более разительные результаты, подтверждающие выводы теории, получил недавно Кондратьев, определив концентрацию атомного водорода, возникающего в разреженном пламени водорода. Воспользовавшись методом термоэлектрического зонда (изучая разогрев термопары, покрытой катализатором для процесса рекомбинации атомов Н), Кондратьев (1945) нашел, что до 30% исходного водорода может переходить в атомное состояние в условиях разреженного пламени. Этот результат нельзя не признать рекордным.

Радикалы ОН были найдены Кондратьевым и сотрудниками (1940—1947) во многих реакциях окисления (окисление влажного СО, ацетилена и др.). В этих же случаях констатировано образование больших количеств атомного водорода. Спектроскопическим методом Кондратьев открыл также образование радикалов CS и SO при окислении CS_2 и COS.

Весьма интересные результаты получил Эмануэль (1939—1942), изучая механизм окисления сероводорода. С помощью метода спектров поглощения ему удалось обнаружить активный промежуточный продукт этой реакции — моноокись серы SO. Было выяснено, что именно SO

определяет целиком скорость процесса. Это, в частности, было констатировано путем добавки к исходной смеси определенных количеств SO, отдельно синтезированной в разряде. В начальной фазе реакции весь окисляющийся H_2S практически целиком переходит в SO, а максимальные количества SO, возникающего в ходе реакции, составляют около 20% от исходного H_2S. Полученные Эмануэлем данные позволяют сформулировать вероятную гипотезу о механизме окисления сероводорода. В самое последнее время еще более интересные результаты были получены Эмануэлем при изучении промежуточных продуктов окисления ацетальдегида. Было показано, что при вымораживании продуктов они становятся нестабильными и при повторном нагревании до температуры реакции самопроизвольно взрывают.

МЕХАНИЗМ РЕАКЦИИ

Конечную задачу химической кинетики можно видеть в том, чтобы, выяснив природу активных радикалов или других промежуточных продуктов, ведущих реакционную цепь, узнать их реакционную способность и, просуммировав все элементарные процессы, получить скорость сложной реакции и все ее закономерности. В настоящее время эта задача уже разрешена для одного очень сложного, но чрезвычайно подробно изученного примера — для реакции водорода с кислородом (Семенов, Кондратьев, Налбандян, Воеводский (1939—1944) у нас; Льюис и Эльбе в США).

Приведем здесь некоторые новые результаты, полученные при изучении этой важнейшей реакции. Налбандян (1946) подробно исследовал обрыв реакционных цепей на различных поверхностях. Достаточно ввести в реакционный сосуд маленькую металлическую проволочку, чтобы резко затруднить протекание реакции. Действие различных поверхностей обнаруживает строгий параллелизм с их способностью вызывать рекомбинацию свободных атомов водорода. Можно даже воспользоваться этим эффектом как способом измерения кинетики рекомбинации. Факт накопления больших концентраций атомного водорода в ходе реакции приводит к выводу, что пределы зажигания и потухания должны сильно отличаться друг от друга. Это предсказание теории было подтверждено экспериментально в работе Ковальского и Медведевой. Налбандян де-

тально исследовал фотохимическую реакцию и подтвердил количественно выводы цепной теории. Единственная гипотеза, не получившая еще окончательного подтверждения, — это наличие радикала HO_2. Доказать на опыте его существование — основная задача в дальнейшем изучении реакции водорода с кислородом.

МАКРОСКОПИЧЕСКАЯ КИНЕТИКА И СВЯЗЬ С ГЕТЕРОГЕННЫМИ РЕАКЦИЯМИ

Кроме элементарного механизма реакции, весьма важна обстановка процесса. Здесь большая заслуга советской школы в подчеркивании и выявлении роли тепловых, диффузионных и гидродинамических факторов при реальном протекании химического процесса в природе или в технике. В развитии этого направления, которое получило у нас название макроскопической кинетики, много сделали Зельдович, Рогинский, Елович, Франк-Каменецкий. Особенный интерес представляет здесь вопрос о связи между гомогенными и гетерогенными реакциями. Принято было считать, что это два совершенно различных класса реакций (первые происходят в объеме, вторые на твердой поверхности) и что никаких переходов между ними нет. Теперь все более выясняется, что очень часто мы имеем дело с тесным переплетением объемных и поверхностных эффектов.

С самого начала было известно громадное значение гетерогенного обрыва реакционных цепей в обычных окислительных цепных реакциях. Особенно наглядно эта роль продемонстрирована в последней работе Налбандяна и Шубиной (1946). Ничтожное изменение поверхности может в сотни раз менять скорость окислительных реакций.

Поляков в течение ряда лет высказывал и качественно доказывал идею о существенной роли поверхности в процессах зарождения и разветвления цепей (для окислительных реакций). Недавно Ковальский (1946) разработал метод раздельного калориметрирования, позволяющий непосредственно измерять объемную и поверхностную реакции. С помощью этого метода Ковальский показал для ряда реакций, которые считали типично гетерогенными, что в действительности они протекают цепным путем в объеме и только инициируются катализатором. Наиболее изученным примером такого рода является реакция восстановления серы из сернистого газа окисью углерода. Очень интересные результаты

получены Ковальским для каталитической реакции водорода с хлором, где добавление кислорода заставляет реакцию локализоваться у поверхности.

Связь между гомогенными и гетерогенными реакциями устанавливается, с другой стороны, открытием нового типа катализа, где реакция протекает в растворе, адсорбированном на поверхности катализатора. Таковы явления ионного катализа в полимолекулярных адсорбционных слоях, предсказанные Семёновым и обнаруженные экспериментально в работах Чиркова и Гольданского (1946). Подобные явления могут иметь довольно большое значение в катализе. Вероятно, таков механизм катализа фосфорной и серной кислотами. Не исключена возможность аналогичных явлений при кислотно-щелочном катализе и в мономолекулярных слоях.

В области катализа работа в основном шла по пути выявления влияния среды и в меньшей степени — по вскрытию механизма самого химического процесса. Здесь, так же как и в реакциях в растворах, не исключена возможность участия радикалов и цепей. Эти разделы химической кинетики развиты у нас пока еще далеко недостаточно. Зато очень широкое развитие получило применение химической кинетики к очень важной в теоретическом и практическом отношении области явлений горения и взрыва.

ТЕОРИЯ ГОРЕНИЯ

Основное содержание работ советской школы в области теории горения состоит в том, что этими работами был построен мост между теорией горения и химической кинетикой.

Ряд исследователей (Вант-Гофф, Ле-Шателье, Таффанель, Нернст) уже давно отдавали себе отчет в том, что горение не что иное, как протекание химической реакции в условиях, благоприятствующих резкому ее ускорению за счет разогрева, что все законы и количественные характеристики явлений горения могут быть сведены к кинетическим, тепловым и диффузионным законам и величинам.

И тем не менее до недавнего времени в науке о горении широким, можно сказать, господствующим распространением пользовались такие величины, как, например, «температура воспламенения». С этой величиной обращались так, как если бы она была первичной, элементарной характеристикой вещества. Ее определяли только экспериментально,

и из нее выводили другие величины, не вдумываясь более глубоко в ее внутренний смысл. В большом ходу были, например, теории, вычислявшие скорость распространения пламени из температуры воспламенения. Таким образом, одну эмпирическую характеристику явления пытались вывести из другой, столь же эмпирической характеристики. Но эти теории совершенно не входили во внутренний механизм явления, в глубине которого всегда скрыта кинетика химического процесса.

После работ советской школы удалось на деле осуществить то, на что классики физической химии могли только намекать.

Чтобы знать все о явлениях горения, мы должны только знать, как выделяется тепло в результате реакции и как это тепло обратно воздействует на реакцию. Если известна кинетика реакции, то мы можем предвычислить условия воспламенения и скорости горения. Если кинетика реакций, происходящих в пламени, не известна, то изучение явлений горения мы можем использовать как средство для изучения этой кинетики, для познания механизма химического процесса.

ВОСПЛАМЕНЕНИЕ

Новое направление в теории горения начинается с исследований Семенова по теории воспламенения (1928). В этих исследованиях было впервые установлено, что существуют два совершенно различных вида воспламенения: цепное и тепловое. Тепловое воспламенение связано с прогрессивным саморазогревом, цепное — с накоплением активных промежуточных продуктов (активных центров). Для цепного воспламенения характерны замечательные явления верхнего и нижнего пределов по давлению: воспламенение можно вызвать как повышением, так и понижением давления. Эти явления были объяснены обрывом реакционных цепей в объеме и на стенках сосуда.

Условие теплового воспламенения определяется возможностью теплового равновесия между реагирующей системой и окружающей средой. Семенов впервые сформулировал это условие математически, применив основной закон химической кинетики: закон Аррениуса о зависимости скорости реакции от температуры. Таким образом, удалось найти связь между давлением и температурой воспламенения. Эта теория была развита математически Тодесом (1933—1940), который подробно рассчитал, как раз-

вивается разогрев во времени в зависимости от кинетики реакции. Экспериментально найденная Семеновым связь между давлением и температурой воспламенения была подтверждена впервые Загулиным, а затем множеством других советских и иностранных авторов. Но в этой теории неопределенными оставались условия теплоотдачи. Поэтому нельзя было рассчитать теоретически абсолютное положение предела воспламенения. Зная температуру воспламенения при одном давлении, можно было рассчитать, как изменится она при другом давлении, но, зная только кинетику реакции и не производя никаких экспериментов по горению, нельзя было еще предсказать, когда же начнется горение.

Эта неопределенность была устранена Франк-Каменецким (1939). Он рассмотрел детально пространственное распределение температур и механизм теплоотдачи при протекании химической реакции с выделением тепла. После этого стало возможно предвычислять условия воспламенения по известной кинетике реакции. Практически еще более важной оказывается обратная задача. Во многих случаях удается использовать измерение условий воспламенения как средство для изучения кинетики и механизма химических реакций, приводящих к воспламенению.

РАСПРОСТРАНЕНИЕ ПЛАМЕНИ

После того как простейшую задачу теории горения, задачу о самовоспламенении, удалось свести к химической кинетике в сочетании с теорией теплопередачи, такая же задача встала по отношению к скорости горения. Поставить теорию распространения пламени на фундамент химической кинетики удалось в основном Зельдовичу и его школе. В работе Зельдовича и Франк-Каменецкого (1938) нормальная скорость теплового распространения пламени была впервые рассчитана, исходя из правильных представлений о кинетике реакции и об условиях ее протекания. Эта теория была подтверждена экспериментально в работах Кокочашвили (1942) по распространению пламени в смесях брома с водородом, Беляева (1940) по горению нитрогликоля, Садовникова и Барского (1941) по горению окиси углерода и углеводородов. В работах Зельдовича и Семенова (1940), исходя из анализа скоростей распространения пламени, удалось выяснить цепной механизм горения окиси углерода при высоких температурах. Далее, Зельдовичем (1941) была разработана теория пре-

делов распространения пламени, позволяющая установить, в каких условиях вообще возможно горение, и объясняющая, в частности, с тепловой точки зрения невозможность прохождения пламени через узкие трубки.

ДЕТОНАЦИЯ

Есть два совершенно различных по скорости вида распространения горения. До сих пор шла речь о нормальном или медленном горении. Но горение может распространяться и с громадными скоростями, быстрее звука, которые измеряются километрами в секунду. Такое горение называется детонационным. Явления детонации очень важны для двигателей и взрывчатых веществ; интересны они и в чисто научном отношении.

Давно научились рассчитывать скорость детонации. В противоположность скорости медленного горения скорость детонации никак не связана с химической кинетикой. Ее можно рассчитать, пользуясь только термодинамикой и газодинамикой. Казалось, что теория детонации не имеет отношения к кинетике химического процесса. Но работами советской школы было показано, что и здесь для полного понимания явления нельзя обойтись без кинетики.

Не всякая горючая смесь может детонировать. Соколик (1934—1938) впервые отметил, что способность детонировать связана со скоростью химических реакций горения, в противоположность скорости распространения детонации, которая от химической кинетики не зависит. Когда к смеси окиси углерода с кислородом добавлялись малые количества водорода, то резко менялась скорость химической реакции горения, а термодинамические свойства смеси практически не изменялись. В соответствии с этим не менялась и скорость детонации, но расширялись те пределы концентрации, в которых возможно распространение детонации. Впоследствии Зельдовичу удалось еще глубже внедрить химическую кинетику в теорию детонации. Он показал, что хотя численное значение скорости детонации и не зависит от скорости химической реакции, но для того, чтобы обосновать самый принцип отбора определенной скорости детонации, нужно учесть условия протекания химической реакции во фронте детонационной волны.

Раньше считалось, что между медленным горением и детонацией существует полный разрыв. Теперь мы знаем, что это не так. Исходя из

весьма плодотворной идеи о влиянии турбулентности, Щелкин (1939—1940) исследовал возникновение и распространение детонации в шероховатых трубах. При этом были открыты совершенно новые явления аномального распространения горения со скоростями, промежуточными между нормальным и детонационным горением. Зельдович (1940) дал теоретическое объяснение этих явлений, связанных либо с шероховатостью стенок, либо с формированием профиля скоростей.

При исследовании детонации вблизи пределов давно было обнаружено замечательное явление вращательного (винтообразного) распространения, получившее название детонационного спина. Многие наблюдали это явление, но никто не понимал, что это такое. Щелкин подробно исследовал явление детонационного спина и дал основную идею его теоретического объяснения, развитую далее им совместно с Зельдовичем. С помощью специальных приемов ему удалось на опыте изменять скорость распространения спиновой детонации.

ГОРЕНИЕ В ДВИГАТЕЛЯХ И ТУРБУЛЕНТНОЕ ГОРЕНИЕ

Технические приложения науки о горении связаны прежде всего с двигателями внутреннего сгорания. На базе цепной теории химических реакций горения Соколиком и его сотрудниками были развиты новые представления о природе детонационных свойств топлив и действия антидетонаторов, а также предложены новые методы борьбы с детонацией в двигателях.

В последнее время все большее значение приобретают реактивные или ракетные двигатели. Развитие этих двигателей ставит новую техническую задачу: осуществить горение при очень больших скоростях газового потока. Решение этой задачи основывается на теории турбулентного горения (Франк-Каменецкий и Минский; Щелкин, 1943—1946). Помещенные на пути газового потока препятствия (экраны) вызывают турбулентное движение газа, приводящее к ускорению горения. Пламя удерживается за такими экранами, цепляется за них. Решение задачи скоростного горения сводится к созданию завихрений, к турбулизации газового потока в камере сгорания.

Как было показано в фундаментальной работе Ландау об устойчивости плоского фронта пламени, помимо искусственно созданной турбулентности

возможна самопроизвольная автотурбулизация пламени. В самое последнее время появилась работа Зельдовича и Розловского, содержащая первые экспериментальные данные, свидетельствующие об эффекте автотурбулизации. Исследование этого явления—одна из интереснейших задач в дальнейшем изучении процессов горения, открывающая новые перспективы для проведения горения в больших объемах и при больших скоростях.

ГЕТЕРОГЕННОЕ ГОРЕНИЕ

До сих пор речь шла о газовых реакциях. В технике очень важны процессы гетерогенного горения. Таково горение твердого топлива.

Здесь также можно связать химическую кинетику с условиями воспламенения и с новым явлением, которого в обычном горении не было,— явлением потухания (Франк-Каменецкий, 1939); воспламенение и потухание тождественно связаны с переходом реакции из диффузионной области в кинетическую и обратно.

Важнейший случай гетерогенного горения — горение угля подробно изучалось рядом советских исследователей (Чуханов и Гроздовский, Предводителев, Блинов, Франк-Каменецкий, Кнорре, Вулис). Здесь была детально изучена роль диффузионных, а в последних работах Франк-Каменецкого также и роль тепловых факторов. Франк-Каменецкому (1943) удалось с помощью исследования явлений воспламенения подойти к изучению кинетики реакции в условиях, где прямое измерение ее затруднительно. Отсюда он пришел к выводу, что реакция имеет не первый, как полагали все предшествующие исследователи, а дробный порядок по кислороду.

Совершенно подобны горению по своему тепловому режиму некоторые каталитические процессы, которые можно назвать каталитическим горением, например окисление аммиака или водорода на платиновых катализаторах (Бубен и Франк-Каменецкий, 1941), окисление спиртов на медных и серебряных катализаторах (Франк-Каменецкий, 1946). Более сложны технические процессы так называемого беспламенного, или каталитического, горения. Для них пока еще не ясно, в какой мере они действительно протекают на поверхности катализатора.

ХОЛОДНЫЕ ПЛАМЕНА

Когда мы говорим о горении, мы обычно связываем это понятие с мощным выделением тепла. Конечно, для всякого технического горения это будет так, здесь задача именно в том и состоит, чтобы получить тепло. Но с точки зрения чисто научной очень интересны так называемые холодные пламена, где мы наблюдаем всю типичную картину явлений горения, а между тем количество выделяющегося тепла может быть совсем небольшим, так что повышение температуры во фронте пламени ничтожно. Распространение холодных пламен связано уже не с передачей тепла, но с диффузией активных промежуточных продуктов цепной реакции. Поэтому изучение холодных пламен представляет особый интерес в связи с цепной теорией кинетики химических реакций.

Давно и хорошо известны холодные пламена углеводородов. Здесь химический процесс окисления не доходит до конца, и вместо продуктов полного сгорания — углекислоты и воды — получаются различные органические вещества. Особенно подробно эти явления изучали: за границей — Ньюитт, Тоуненд, Эгертон; у нас — Нейман и его сотрудники (1936—1939). Они приписывают активную роль в распространении холодного пламени органическим перекисям. Но в этих холодных пламенах все же имеет место заметный разогрев.

Гораздо более замечателен исследованный Воронковым и Семеновым (1939) случай холодного пламени сероуглерода. Авторы работали с крайне бедными смесями (сотые доли процента CS_2), для которых температура во фронте пламени практически не отличается от температуры исходной газовой смеси. Пламя распространяется по газу, температура которого может быть на 100° ниже температуры воспламенения, невзирая на то, что во фронте пламени смесь нисколько не нагрета. Это совершенно парадоксальное и необычайное явление остается пока единственным в своем роде, если не считать некоторых аналогий в биологических процессах. Объяснения его можно искать только в круге идей цепной теории, в идее взаимопомощи цепей. В зоне реакции, где концентрация активных центров велика, это повышение концентрации способствует, очевидно, разветвлению цепей.

ПОЛУЧЕНИЕ ВЫСОКИХ ТЕМПЕРАТУР И ДАВЛЕНИЙ МЕТОДОМ АДИАБАТИЧЕСКОГО СЖАТИЯ

Харитоном, Клязером, Рейновым и Рябининым осуществлена установка, дающая возможность путем быстрого адиабатического сжатия газа получать одновременно давления до 10 000 кг/см² и температуры до 3000°, т. е. фактически получать состояния, приближающиеся к тем, которые имеют место при детонации взрывчатых веществ.

Из интересных результатов, полученных с помощью этой установки, следует отметить образование окислов азота в количестве около 0.5% после однократного адиабатического сжатия воздуха.

ВЗРЫВЧАТЫЕ ВЕЩЕСТВА

Большая группа работ проводится советскими учеными в области изучения механизма детонации, действия взрыва, горения и кинетики медленного разложения взрывчатых веществ (в. в.). Рассмотрим результаты, полученные в основном школой Института химической физики под руководством Харитона.

К и н е т и к а м е д л е н н о г о р а з л о ж е н и я в. в. Изучение кинетики медленного разложения в. в., начатое Рогинским с сотрудниками (1929—1930), установило некоторые особенности кинетики разложения вторичных в. в. (большие значения энергии активации и предэкспоненциального множителя в выражении для скорости реакции), выявило чрезвычайно важную универсальную зависимость, согласно которой логарифмы задержки вспышки связаны с обратными температурами, и показало ту исключительно важную роль, которую играет в кинетике разложения в. в. автокатализ. В случае твердых в. в. важнейшую роль играют топохимические факторы.

В дальнейшем Андреевым было проведено изучение разложения некоторых азидов и изучена для большого числа веществ связь периода задержки вспышки с температурой.

Апиным была изучена кинетика разложения ряда нитроэфиров, азида свинца. Особенно интересные результаты были им получены при изучении кинетики разложения сверхчувствительного хлористого азота.

Было установлено наличие верхнего предела вспышки для паров хлористого азота. При откачке системы, в которой находился хлористый

азот при комнатной температуре, по достижении давления в 40 мм рт. столба происходила самопроизвольная вспышка, переходящая во взрыв при контакте с жидким хлористым азотом. Самопроизвольная вспышка чистых паров происходит также при любых давлениях, вплоть до $2 \cdot 10^{-3}$ мм. Поэтому хлористый азот, замороженный жидким воздухом, при оттаивании (при отсутствии воздуха) дает взрыв.

Перед войной Боболевым было начато изучение кинетики и некоторых свойств сверхчувствительной азотисто-водородной кислоты, а также хлористого азота. Было установлено очень интересное явление, заключающееся в том, что реакция протекает в отдельных точках на поверхности сосуда, причем наблюдается интенсивное свечение этих точек.

В самое последнее время Рябинин провел разведочные опыты по изучению действия статического сверхвысокого, до 50 000 кг/см^2, давления на скорость разложения в. в. Установлено, что такое давление в очень малой степени изменяет скорость разложения в. в.

Механизм горения в. в. и переход горения в детонацию. Существенные результаты были получены Беляевым по вопросу о механизме горения в. в. Очень плодотворной оказалась выдвинутая и обоснованная им гипотеза о решающей роли реакции в газовой фазе при горении в. в. Прямыми опытами доказано наличие темного промежутка между пламенем и поверхностью вещества. В дальнейшем Зельдовичем и его сотрудниками эта гипотеза была распространена на нелетучие бездымные пороха. Было предположено, что их первичное разложение дает газовые продукты, при дальнейшей реакции которых в газовой фазе выделяется основная теплота реакции; скорость горения пороха определяется именно этой реакцией в газовой фазе. Беляевым была впервые успешно применена уточненная тепловая теория Зельдовича — Франк-Каменецкого и на примере нитрогликоля показано соответствие между вычисленным и измеренным значениями скорости горения. Это предположение получило существенную поддержку опытами Лейпунского и Аристовой, которые показали, что при горении пороха основное тепло выделяется в газовой фазе, а максимальная температура на самой поверхности горящего пороха составляет всего лишь 300° С при температуре горения в 2000—3000° С.

Зельдовичем с сотрудниками был решен ряд очень важных вопросов внутренней баллистики, связанных с горением пороха, с устойчивостью

этого горения, с влиянием на горение газодинамических факторов и т. д.

Беляевым создана методика изучения горения инициирующих взрывчатых веществ и начато широкое изучение этой, ранее совершенно неисследованной области. Оказалось, что ряд инициирующих в. в. способен при определенных условиях к спокойному горению.

При экспериментальном исследовании горения инициирующих в. в. (способных к горению), проведенном Беляевым и Беляевой, были получены очень интересные факты (например, беспламенное горение гремучей ртути) и установлены новые важные закономерности, в частности близкая к линейной зависимость скорости горения от давления.

Беляевым разработана оригинальная теория перехода горения в детонацию на основе газодинамической схемы, связывающей возникновение детонации с достижением критической скорости горения и с нарушением газового баланса. Эта схема в настоящее время является наиболее достоверной и обоснованной как теоретически, так и надлежащими экспериментальными исследованиями.

В о з н и к н о в е н и е д е т о н а ц и и. Широко изучалось возникновение детонации в. в. под действием разнообразных агентов. В этих опытах в. в. подвергались действию ионной и электронной бомбардировки, бомбардировки α-частицами, а также действию микроскопических частиц, летящих с большой скоростью. Особенно детально было изучено возникновение детонации при действии удара и микроскопических тепловых разогревов значительной интенсивности.

Соответствующие опыты проводились Харитоном, Беляевым, Рдултовской.

Результаты всех этих исследований показали неправильность и полную необоснованность распространенной в иностранной литературе точки зрения, согласно которой для развития детонации достаточно одновременного разложения нескольких соседних молекул. Было установлено, что в действительности развитие детонации возможно лишь при возникновении локального разогрева, охватывающего по меньшей мере миллиарды молекул, хотя в общепринятых единицах измерения эти разогревы микроскопические.

Харитоном была разработана для простейшего случая жидкого в. в. теория тепловых явлений при ударе (к сожалению, эта работа была прервана войной и не закончена), связывающая развитие детонации

при ударе с возникновением локальных микроскопических разогревов. Эта концепция Харитона была подкреплена опытами Беляева, который показал, что локальные разогревы значительной интенсивности, получаемые посредством пережигания проволочек током, вызывают развитие детонации при условии, что эти локальные разогревы возникают при наличии высокого давления. Отметим, что в момент удара высокое давление налицо.

ЭКСПЕРИМЕНТАЛЬНАЯ МЕТОДИКА ИЗУЧЕНИЯ ВЗРЫВОВ

Для развития теоретических представлений исключительно важной была работа по усовершенствованию экспериментальных методов изучения взрыва. В практике широко распространены различные «пробы» для испытания взрывчатых веществ (Трауцля, Гесса, Каста). В отношении ряда этих «проб» до последнего времени было неясно, что же они собственно измеряют, в каких условиях их можно применять.

Исследованиями Садовского и Похила, Харитона, Беляева были глубоко проанализированы работа и действие заряда в условиях этих методов испытания.

Тонкие физические методы, в частности осциллография, были применены к исследованию взрыва в работах Степанова.

В фундаментальных хорошо известных работах Садовского широко исследовался ряд вопросов, связанных с действием взрыва: вопросы о безопасных расстояниях, о разрушительном действии ударной волны, о связи импульса взрыва с величиной заряда и расстоянием, о местном действии взрыва, о структуре кумулятивного пучка и т. д.

Соболев, Аленцев, Беляев и Степанов изучали оптическим методом температуру детонации нитроглицерина, нитрогликоля и метилнитрата.

ДЕТОНАЦИОННАЯ СПОСОБНОСТЬ И ВЛИЯНИЕ ФИЗИЧЕСКИХ ФАКТОРОВ

Основной стержневой проблемой, изучавшейся за последнее десятилетие, была проблема детонационной способности, устойчивости детонации и влияния физических факторов на устойчивость и эффект детонации.

Важнейшим в развитии этой проблемы было установление Харитоном руководящего теоретического принципа, роль и значение которого в последующей разработке теории была исключительно важна.

Исходя из того, что химическое превращение во фронте детонации не является мгновенным, Харитон утверждает, что для в. в. устойчивая детонация возможна лишь в том случае, если за время реакции во фронте детонации не успевает произойти разброс вещества, производимый развивающимися при взрыве колоссальными давлениями.

Из теории Харитона вытекает очень важная роль диаметра заряда: при достаточно большом диаметре могут устойчиво детонировать даже вещества с очень малой скоростью реакции взрывчатого превращения. Вывод этот был подтвержден Беляевым и Харитоном опытами с аммиачной селитрой. Помимо изложенных теоретических идей, по этой проблеме было получено много оригинального экспериментального материала.

Беляевым, Рдултовской и Беляевой было проведено обширное исследование влияния физических факторов на устойчивость детонации аммиачно-селитренных в. в. Количественно изучено влияние степени измельчения. Изучение влияния плотности привело к новым представлениям в вопросе о предельной плотности. Изучено влияние оболочки, и в первом приближении установлен механизм ее действия, изучено влияние мощности инициирования. Установлены некоторые характерные явления, присущие детонации вблизи предельного диаметра.

Аналогичные исследования с однородными взрывчатыми веществами (тротил, пикриновая кислота, тэн, гексоген) были проведены Боболевым. Им был открыт ряд оригинальных фактов и, в частности, установлено сильное влияние степени измельчения на детонационную способность однородных в. в. Было установлено, что в случае однородных в. в., в противоположность смесям, увеличение плотности повышает устойчивость детонации.

Существенные результаты были получены Харитоном и Ратнером при изучении детонации жидких в. в. (нитроэфиры). Было установлено, что в нитроэфирах при достаточно малом диаметре заряда большая скорость (7—8 км/сек) скачком переходит в малую (2 км/сек).

К перечисленным выше работам непосредственно примыкает работа по важнейшему вопросу об уравнении состояния газов. Развивая фундаментальную работу Ландау и Станюковича, предложивших на основе

анализа зависимостей скорости детонации от плотности новое уравнение состояния газов взрыва при высоких плотностях, Харитон указал, что температура детонации должна зависеть от плотности в. в. Ратнером был проведен ряд расчетов по вопросу о состоянии вещества во фронте детонации.

ВЗРЫВНОЕ ГОРЕНИЕ

Важным теоретическим достижением является введение Апиным и Харитоном нового представления о механизме детонации. Обычно предполагалось, что реакция во фронте детонационной волны обусловлена разогревом, вызываемым сжатием. Согласно новому представлению, реакция вызывается поджиганием струями газов с последующим сгоранием при крайне высоких давлениях, характерных для детонации («взрывное горение»).

Есть все основания полагать, что взрывное горение является более универсальным и распространенным механизмом взрывного разложения не только в случае смесей (для которых эта гипотеза первоначально была предложена), но и в случае однородных в. в.

КИНЕТИКА — ВАЖНЫЙ РАЗДЕЛ ХИМИЧЕСКОЙ НАУКИ

Химическая кинетика — не только глава физической химии. На наших глазах она превращается в один из основных разделов общей химии. Она охватывает все приложения химии — от химической промышленности до взрывчатых веществ и биологии. Именно здесь, в этом широком развитии химической кинетики, в последние годы особенно большую роль сыграла советская наука. В настоящем обзоре мы подробно остановились на работах только одной школы. Но создание советской химической кинетики — дело целого ряда школ. Здесь сыграли свою роль и фотохимики (Теренин, Кондратьев и др.), и биохимики (Парнас, Опарин, Энгельгардт и др.), и школа Института имени Карпова (Сыркин, Темкин, Медведев и др.), и электрохимики (Фрумкин), и каталитики (Зелинский, Баландин, Рогинский, Кобозев и др.), и пирогенетики (Фрост, Васильев и др.), и представители украинской науки (Бродский, Поляков), и ряд других исследователей (Шилов, Ерофеев и др.).

Хотя химическая кинетика и очень молодая наука, все же советские ученые успели занять здесь почетное место.

Академик
А. Н. Фрумкин

ФИЗИКО-ХИМИЧЕСКИЕ ЯВЛЕНИЯ ПРИ ДВИЖЕНИИ НА ГРАНИЦЕ ДВУХ ЖИДКОСТЕЙ

Современное развитие химии характеризуется все большим проникновением физических представлений в область химических фактов и все более широким использованием в химии физических методов исследования. Это позволило в ряде случаев влить новое содержание в важные разделы химической науки. Наряду с этим происходит, однако, и обратный процесс «химизации» ряда разделов классической физики. При более глубоком рассмотрении многих явлений, которые обычно трактовались без учета физико-химических особенностей вещества, оказывается, что в действительности физико-химические факторы играют в них существенную, а иногда и решающую роль. Это относится, в частности, к явлениям деформации твердых тел и течения жидкостей. Исследование механических свойств реальных систем, особенно развившееся в связи с широким использованием в технике новых искусственных материалов (пластмассы, каучуки, искусственное волокно и др.), показало недостаточность того аппарата, которым пользовалась классическая механика идеально упругого твердого тела и ньютоновской жидкости.

Ряд исследователей, следуя американскому ученому Бингэму, искали выход из этого положения в создании новой науки — «реологии» (от ρεο — течь), науки о механических свойствах тел, промежуточных между твердыми телами и жидкостями и обладающих свойством пластичности,

т. е. способностью течь под влиянием внешних сил. Однако развитие реологии имело по преимуществу формальный характер и почти не привело к созданию новых физических представлений. Больший успех имели, как мне кажется, те работы, в которых делались попытки непосредственно применить физико-химические представления к объяснению механических свойств. Особенно существенные результаты были получены у нас и за границей (Марк, Кун, Эстбери, Френкель, Каргин, Александров) в области высокомолекулярных соединений. Применение методов статистической физики и теории строения вещества позволило далеко продвинуться в объяснении упругих и пластических свойств высокополимеров и связать макроскопическое механическое поведение этих тел со свойствами их отдельных молекул. Наряду с изучением высокополимерных соединений такое направление, которое можно было бы назвать физико-химической механикой, привело к ряду важных результатов и в других случаях. Достаточно упомянуть, например, о работах академика Ребиндера, показавших, какое значение при процессах деформации и разрушения твердых тел имеет проникновение в микрощели на поверхности тела адсорбирующихся веществ, и выявивших огромное влияние внешней среды на прочность и твердость.

В настоящей статье приведены основные разултаты исследований, проведенных в Институте физической химии Академии Наук СССР и посвященных выяснению роли физико-химических явлений в одном из простейших случаев движения, а именно, движения на границе двух жидкостей или на границе между жидкостью и газом.[1] Интерес к этой проблеме существовал давно. Он был связан с известным уже в древности явлением гашения волн пленкой масла, которое упоминается и в стихах Тютчева:

И на бунтующее море
Льет примирительный елей.

Рамки настоящей статьи не позволяют мне подробнее остановиться на теории этого интересного явления, в выяснении механизма которого выдающуюся роль сыграли работы советских исследователей [1, 2], хотя теория эта имеет много точек соприкосновения с предметом настоящей статьи. Замечу только, что использованный нами способ трактовки влияния

[1] Теоретическая часть этой работы проводилась В. Г. Левичем и автором, экспериментальная — Т. А. Крюковой и И. А. Багоцкой.

физико-химических факторов на условия движения на жидких границах раздела в некоторых отношениях существенно отличается от способа, использованного Р. Ивановым и В. Шулейкиным, и более близок к способу, применявшемуся В. Левичем.

При рассмотрении движения жидкости относительно твердого тела в гидродинамике принимается, что касательная скорость жидкости у поверхности твердого тела равна нулю, т. е. что отсутствует скольжение. Это условие можно рассматривать как условие физико-химического характера. Однако для его практического выполнения достаточно, чтобы трение между слоем жидкости, прилегающим к твердому телу, и его поверхностью было того же порядка, что и трение между двумя слоями жидкости. При этих условиях скорость жидкости непосредственно у поверхности твердого тела будет мала по сравнению со скоростью на макроскопических расстояниях от него, и ею можно будет пренебречь.[1]

В случае границы между двумя жидкостями требование равенства касательной скорости нулю отпадает и заменяется условием равенства касательных скоростей обеих жидкостей у общей границы раздела. Если бы дело этим ограничивалось, то тангенциальное движение у поверхности раздела могло бы происходить беспрепятственно, и для случая, когда вязкости обеих жидкостей — величины одного и того же порядка, условия движения на границе раздела в этом смысле не отличались бы от условий в середине каждой из фаз.

Более подробное рассмотрение показывает, однако, что вывод этот правилен только при определенных свойствах поверхности жидкости, а именно: пограничное натяжение на этой поверхности не должно зависеть от состояния движения; натяжение поверхности, которая подвергается сжатию или расширению, не должно отличаться от натяжения покоящейся поверхности. Такое условие выполняется для чистой поверхности индивидуальной жидкости, по крайней мере при не слишком больших скоростях деформации. Возможно, что в случае очень больших ско-

[1] Условия движения жидкости вблизи поверхности твердого тела усложняются при наличии силового поля, непосредственно действующего на слой жидкости, прилегающий к поверхности, например в случае электроосмотического движения жидкости. Однако скорости, которые при этом возникают, по порядку величины малы по сравнению со скоростями у жидких поверхностей раздела, и я на них дальше не буду останавливаться.

ростей строение поверхностного слоя может отличаться от равновесного, например из-за разницы в степени ориентации молекул на поверхности; однако достаточно надежных данных по этому вопросу не имеется (см. ниже).

Иначе обстоит дело с поверхностью, на которой имеется какой-нибудь адсорбированный слой, например ионов или молекул растворенного вещества. В этом случае растяжение и сжатие поверхности приводят к изменению поверхностной концентрации адсорбированных частиц, а следовательно и к изменению пограничного натяжения. Последнее будет возрастать при увеличении концентрации адсорбированного вещества и уменьшаться в обратном случае; таким образом, изменение пограничного натяжения всегда приводит к возникновению сил, противодействующих деформации поверхности и затрудняющих движение.

Зависимость пограничного натяжения от деформации поверхности в статических условиях была впервые рассмотрена Гиббсом [3]. Гиббс ввел представление о модуле упругости поверхности, выражающем изменение пограничного натяжения σ при изменении величины поверхности s на ds. Модуль упругости поверхности A можно, следовательно, определить соотношением

$$d\sigma = A\frac{ds}{s}. \tag{1}$$

При этом предполагается, что при изменении величины поверхности сохраняется равновесие между поверхностным слоем и непосредственно прилегающим к нему слоем жидкости. Здесь я не останавливаюсь подробнее на предложенном Гиббсом способе рассмотрения, так как последний относится к случаю, отличному от того, который нас будет интересовать в дальнейшем; однако, следуя ходу мыслей Гиббса, не трудно найти значения модуля упругости для различных состояний пограничного слоя. Таким образом, получаются следующие выражения для A:

1) На поверхности находится двойной электрический слой, причем плотность заряда равна ε, а емкость двойного слоя C

$$A = \frac{\varepsilon^2}{C}. \tag{2}$$

2) Поверхность адсорбирует молекулы растворенного вещества из объемного слоя толщиной h, причем концентрация адсорбируемого ве-

33*

щества равна c, а равновесное значение адсорбции на единицу поверхности Γ_0

$$A = \frac{RT\Gamma_0^2}{ch}, \qquad (3)$$

R обозначает здесь газовую постоянную, а T — температуру.

3) На поверхности находится слой молекул, которые не могут переходить в объем; их поверхностная концентрация равна Γ_0. В этом простейшем случае

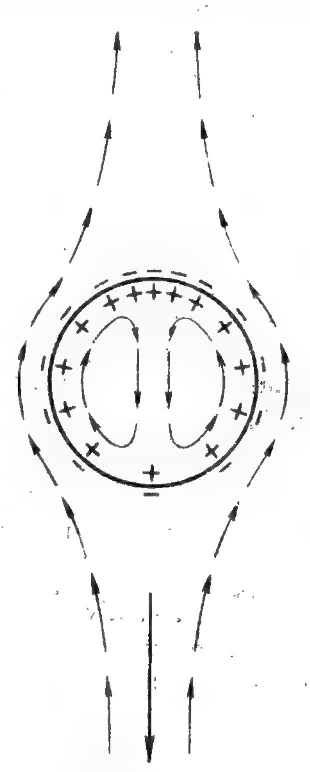
$$A = \Gamma_0 \left| \frac{\partial\sigma}{\partial\Gamma_0} \right|. \qquad (4)$$

Представления Гиббса об изменчивости пограничного натяжения в присутствии адсорбированных молекул часто использовались качественно, например для объяснения устойчивости жидких пленок (так называемый эффект Гиббса — Марангони); однако они не подвергались количественной проверке. Осталось без внимания и то обстоятельство, что модуль упругости в случае адсорбции растворимых молекул оказывается пропорциональным квадрату количества адсорбированного вещества, в то время как обычно действие адсорбированного вещества сказывается лишь в первой степени.

Причина того, что представления Гиббса были мало использованы, заключается в следующем. В реальных условиях, особенно при наличии заряженных ионов или адсорбированных молекул, мы имеем дело не с зависимостью пограничного натяжения от величины деформации, а с его зависимостью от скорости деформации. Иначе говоря, требуется решение задачи не в статических, а в динамических условиях. Дело в том, что наряду с переносом адсорбированных частиц движущейся поверхностью, приводящим к нарушению равномерности распределения этих частиц по поверхности и к изменению пограничного натяжения, всегда существуют процессы рассеяния, вызывающие выравнивание этих нарушений.

Поясним это на конкретном примере движения заряженной поверхности. Рассмотрим те движения, которые возникают при падении жидкой капли или поднятии пузырька газа в жидкой среде. Картина этих движений (в разрезе) изображена на фиг. 1, где скорость движения жидкости (по отношению к системе координат, связанной с каплей) обозначена стрелками. Падающая капля омывается окружающей ее жидкостью.

При этом в капле возникают вихреобразные движения: вблизи поверхности капли жидкость в ней движется вместе с окружающей средой, а в середине капли жидкость движется в обратном направлении. Легко понять, что поверхность капли в нижней ее части подвергается непрерывному растяжению, а в верхней — сжатию. Если на поверхности находится двойной слой из зарядов противоположного знака, как это схематически изображено на фиг. 1, то плотность зарядов в нижней части капли уменьшается вследствие растяжения поверхности, а в верхней части увеличивается вследствие сжатия. Благодаря этому в окружающей среде появляется электрическое поле и течет электрический ток, который частично выравнивает возникшие различия в плотности зарядов. Сила этого тока, очевидно, зависит от геометрических условий и электропроводности среды. Стационарное состояние системы определяется соотношением между скоростью переноса зарядов движущейся поверхностью и скоростью, с которой происходит выравнивание возникающей разности потенциалов электрическим током.



Фиг. 1. Движение жидкости при падении капли и вызываемое им изменение распределения зарядов двойного слоя.

Маленькие стрелки обозначают направление движения жидкости, большая—направление движения капли в целом

Подобным же образом можно рассмотреть случай, когда на поверхности капли находятся не ионы, а нейтральные молекулы, адсорбированные из раствора. Поверхностная их плотность в нижней части капли из-за постоянного растяжения поверхности будет, очевидно, меньше, чем это соответствовало бы равновесию с исходным раствором; в верхней же части значение поверхностной плотности будет превышать равновесное. Поэтому на нижней части капли все время идет процесс адсорбции из раствора, а в верхней — обратный процесс десорбции. Процессы эти вызывают изменение концентрации адсорбирующегося вещества в прилегающих к

капле слоях раствора; последняя, очевидно, будет уменьшаться вблизи нижнего конца капли и увеличиваться вблизи верхнего. Возникающие разности концентраций выравниваются благодаря объемной диффузии в жидкости, окружающей каплю; скорость этого процесса зависит от коэффициента диффузии растворенного вещества и размешивания жидкости, вызванного движением поверхности. Стационарное состояние, которое устанавливается при этом, определяется, как и в выше рассмотренном случае, взаимоотношением скоростей конвективного переноса вещества движущейся поверхностью жидкости и выравнивания нарушений равновесного состояния системы.

Наконец, можно рассмотреть и случай, когда на поверхности имеется адсорбированный слой нерастворимых молекул. Здесь также нужно учесть выравнивание вызванных конвективным переносом изменений поверхностных концентраций благодаря поверхностной диффузии адсорбированных молекул в направлении, обратном движению поверхности. В. Г. Левичем и автором были даны математические выражения, определяющие зависимость пограничного натяжения от скорости деформации поверхности в условиях падения капли [4]. Обозначим скорость тангенциального движения поверхности через $\vec{v}_{tg}$. Тогда скорость деформации поверхности выражается математически через поверхностную дивергенцию $\vec{v}_{tg}$, т. е. $\operatorname{div}_s \vec{v}_{tg}$. Введем величину e, которую мы определим как отношение изменения пограничного натяжения к скорости деформации поверхности:

$$\Delta\sigma = e \operatorname{div}_s \vec{v}_{tg}. \quad (5)$$

Величину e, в отличие от модуля поверхностной упругости Гиббса A, можно назвать динамической упругостью.

В литературе, например, в работах французского физика Буссинеска, уже использовалось представление о существовании величины, определяемой сходным образом; для нее предлагалось название — «поверхностная вязкость». Однако эти представления существенно отличны от наших. Буссинеск рассматривал поверхностную вязкость как специфическое свойство поверхности жидкости. Между тем, как будет показано дальше, величина e определяется не только свойствами поверхности, но и свойствами системы в целом, в частности природой объемной фазы и

геометрическими размерами. Сам термин «поверхностная вязкость» может привести к недоразумениям, так как под поверхностной вязкостью в литературе по поверхностным явлениям часто понимается другая величина, с которой приходится иметь дело при исследовании течения мономолекулярных слоев в двухмерных капиллярах. На возможность измерения этой величины, аналогичной обычной объемной вязкости, впервые указали Талмуд и Бреслер [5]. Поверхностная же вязкость Буссинеска аналогична не обычной объемной вязкости, а величине, которую в гидродинамике называют «второй вязкостью».

Развитая нами теория приводит к следующим выражениям для e:

1) На поверхности находится двойной электрический слой

$$e=\frac{1}{2}\,\frac{\varepsilon^2}{\varkappa}\,a, \tag{6}$$

где $\varkappa$ — электропроводность раствора, окружающего каплю, a — радиус капли.

2) На поверхности находятся молекулы вещества, адсорбированные из раствора

$$e=\frac{RT\Gamma_0^2}{c}\,\frac{\delta}{D}, \tag{7}$$

где D — коэффициент диффузии растворенного вещества, а δ — толщина так называемого диффузионного слоя Нернста. Величину δ можно определить следующим образом. Если концентрация растворенного вещества у поверхности равна c', а в середине раствора c, то скорость диффузии к единице поверхности равна $\frac{(c-c')\,D}{\delta}$. Величина δ зависит от скорости движения раствора [6] и, как показала последняя работа Левича и Меймана, даже при заданной скорости и данных геометрических условиях является только приближенно постоянной. Уравнение (7) выведено в предположении, что скорость процессов адсорбции и десорбции велика по сравнению со скоростью объемной диффузии.

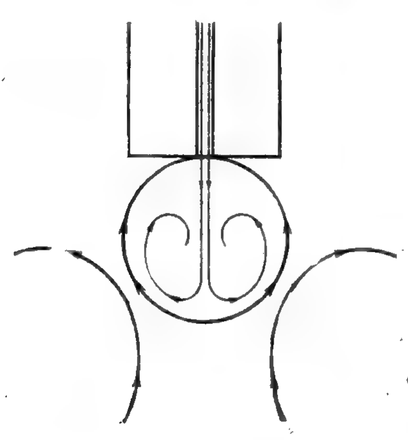
3) На поверхности находятся нерастворимые адсорбированные молекулы

$$e=\left|\frac{\partial\sigma}{\partial\Gamma_0}\right|\frac{\Gamma_0 a^2}{D_s}, \tag{8}$$

где $\left|\frac{\partial\sigma}{\partial\Gamma_0}\right|$, как и в уравнении (4), обозначает равновесное значение произ-

водной от пограничного натяжения по адсорбированному количеству с положительным знаком, а D_s — коэффициент поверхностной диффузии адсорбированных молекул.

Если сопоставить полученные нами выражения для e со значениями модуля поверхностной упругости Гиббса, то оказывается, что отношение $e : A$ имеет всегда размерность времени. Легко убедиться, что это отношение представляет собой время релаксации, необходимое для того, чтобы вызванное движением поверхности изменение равномерного распределения ионов или молекул на поверхности исчезло бы (точнее, уменьшилось бы до $^1/_e$ от своей первоначальной величины). Изложенного достаточно, чтобы пояснить общие физические основы рассмотренной нами группы явлений.



Фиг. 2. Движение жидкости при вытекании капли.

Посмотрим теперь, как сказывается вызываемое движением поверхности изменение ее состояния на наблюдаемых на опыте свойствах системы. При этом используем два примера, а именно: падающую каплю как простейший случай, для которого расчеты могут быть проведены с наибольшей точностью, и каплю, вытекающую из капилляра.

При вытекании капли одной жидкости в другую, например капли ртути в воду, на границе между обеими жидкостями возникают движения, аналогичные движениям вблизи поверхности падающей капли. Движения эти изображены схематически на фиг. 2. Теория, развитая для падающей капли, может быть приближенно перенесена и на вытекающую каплю. Гидродинамические условия в этом случае гораздо сложнее; однако, не говоря уже о том практическом значении, которое имеет исследование явлений на вытекающей капле из-за широкого использования капельного электрода в полярографическом анализе, изучение процессов на вытекающей капле имеет ряд преимуществ, которые будут ясны из дальнейшего. В общем экспериментальное исследование этих явлений и проверка теории могут итти по следующим четырем направлениям:

1) *Прямое измерение скорости движения поверхности.*

Как было разъяснено выше, изменение пограничного натяжения вызывает торможение тангенциального движения на границе раздела. Математическая теория приводит к следующему соотношению:

$$v = v_0 \frac{2\mu + 3\mu'}{2\mu + 3\mu' + \gamma}, \tag{9}$$

где v_0 — скорость тангенциального движения в отсутствие торможения, т. е. для «чистой» поверхности, v — скорость в тех же условиях, но при наличии эффектов торможения, μ и μ' — вязкости окружающей среды и капли соответственно, а γ — величина, определяющая эффект торможения и связанная с ранее рассмотренной величиной e простым соотношением

$$\gamma = \frac{2e}{u}. \tag{10}$$

Проверка уравнения (9) требует прямых измерений скорости движения жидкости вблизи поверхности капли. Такие измерения представляют значительные трудности и не очень точны. Для случая торможения зарядами двойного электрического слоя уравнения (6) и (10) дают $\gamma = \frac{\varepsilon^2}{\varkappa}$, следовательно [7]

$$v = v_0 \frac{2\mu + 3\mu'}{2\mu + 3\mu' + \frac{\varepsilon^2}{\varkappa}}. \tag{11}$$

Как видно из уравнения (11), v делается малым при больших значениях заряда капли ε или малых электропроводностях $\varkappa$.

Уравнение (11) было подвергнуто проверке Т. Крюковой для случая капель ртути, вытекающих в растворы хлористого калия различной концентрации. Величина заряда ε изменялась при этом поляризацией капли, что легко осуществимо в случае капли ртути, вытекающей из капилляра. Величина v определялась визуально по движению взвешенных в жидкости частиц угля, которые увлекаются движущейся жидкостью. В полном соответствии с уравнением (11) оказалось, что тангенциальное движение жидкости происходит либо при очень малых значениях ε, либо при больших $\varkappa$, т. е. в хорошо проводящих растворах. В других случаях скорость v делается очень малой и ею можно пренебречь, иначе говоря, вытекающая капля растет только радиально, как раздувающийся резиновый шарик.

Этот вывод имеет существенное значение для полярографии. Дело в том, что основное количественное соотношение полярографического анализа, связывающее силу тока, который может течь на капельный электрод, с концентрацией растворенного вещества, так называемое уравнение Ильковича — Райдиля — Мак Гиллаври, которое во многих случаях хорошо оправдывается на опыте, выводится в предположении именно такого радиального роста. В практической полярографии условия действительно обычно выбираются таким образом, чтобы обеспечить отсутствие тангенциальных движений (малая электропроводность раствора, присутствие адсорбирующихся веществ). При дальнейшем развитии полярографии начали, однако, во многих случаях пользоваться и хорошо проводящими растворами, в которых, согласно уравнению (11), тангенциальные движения могут беспрепятственно развиваться, и встретились с резкими отклонениями от обычных соотношений. До того, как нами была развита теория тангенциальных движений, отклонения эти в работах, например, американских авторов получили неправильные объяснения, иногда противоречащие основным электрохимическим представлениям [8].

2) *Измерение скорости падения капли.*

Еще Релей обратил внимание на то обстоятельство, что существование тангенциального движения поверхности должно влиять на скорость падения капель. При наличии тангенциального движения вязкое сопротивление среды уменьшается по сравнению с тем сопротивлением, которое встретило бы твердое тело. Математическая теория падения капель была дана Адамаром и Рыбчинским. Они показали, что скорость падения жидкой капли u определяется уравнением

$$u = u_s \frac{3\mu + 3\mu'}{2\mu + 3\mu'}, \tag{12}$$

где u_s — скорость падения твердого шарика того же радиуса, вычисленная по известной формуле Стокса. Если μ значительно больше μ', как, например, в случае падения капли ртути в глицерине или поднятия пузырька воздуха в воде, уравнение (12) приобретает следующий вид:

$$u = \frac{3}{2} u_s, \tag{13}$$

т. е. скорость падения жидкой капли в полтора раза превышает скорость, вычисленную по Стоксу. Так как вопрос о скорости падения капель пред-

ставляет интерес и с точки зрения практической вискозиметрии, то уравнения (12) и (13) неоднократно подвергались проверке, однако в общем с неудовлетворительными результатами. Наблюдаемые скорости падения обычно оказывались гораздо ближе к стоксовым скоростям, чем это следовало из теории. Хотя неоднократно высказывалось предположение, что эти отклонения от теории вызваны какими-то поверхностными явлениями, но до последнего времени не удалось дать удовлетворительного объяснения этого комплекса явлений. Нами было показано [4, 7], что при наличии торможения тангенциального движения скорость падения выражается соотношением

$$u = u_s \frac{3\mu + \gamma}{2\mu + \gamma} \tag{14}$$

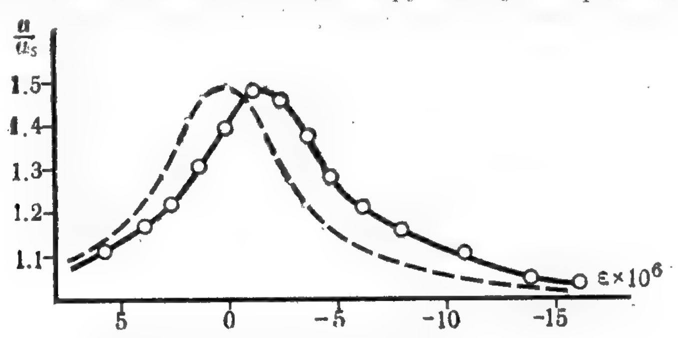
(для простоты ограничиваюсь здесь случаем, когда $\mu \gg \mu'$). Если торможение зависит от присутствия двойного слоя, то $\gamma = \frac{\varepsilon^2}{\varkappa}$ и

$$u = u_s \frac{3\mu + \frac{\varepsilon^2}{\varkappa}}{2\mu + \frac{\varepsilon^2}{\varkappa}}. \tag{15}$$

Таким образом, скорость падения капли должна зависеть от ее заряда и электропроводности среды. Этот вывод нашел свое подтверждение в работе Багоцкой и Фрумкина [9], исследовавших зависимость скорости падения капель ртути в глицерине от величины первоначально сообщенного им заряда. Результаты одной из серий опытов приведены на фиг. 3. Как видно из рисунка, изменение величины заряда действительно позволяет варьировать скорость падения от значений, в полтора раза превышающих скорость по Стоксу (при малых зарядах), до значений, соответствующих скорости падения твердого шарика (при больших зарядах). Опыт подтверждает выводы теории и количественно; небольшое смещение экспериментальной кривой по сравнению с теоретической вызывается, как было показано в цитированной работе, действием следов кислорода, растворенного в глицерине, которые не удается полностью удалить; они несколько смещают величину заряда капли во время ее падения в положительную сторону. Изменение скорости падения, как показали опыты Багоцкой, происходит и при адсорбции растворенных в глицерине поверхностно активных веществ.

3) *Непосредственное доказательство переноса зарядов или адсорбированных частиц движущейся поверхностью.*

Как было показано выше, тангенциальное движение очищает нижнюю часть поверхности и собирает заряды двойного слоя или адсорбированные молекулы в ее верхней части. Это приводит в случае наличия двойного слоя на поверхности к возникновению разности потенциалов между верхним и нижним концом капли, которую и нужно рассматривать как



Фиг. 3. Зависимость скорости падения капли ртути от ее заряда ε в насыщенном растворе Na_2SO_4 в глицерине; $\varkappa = 8.1 \cdot 10^{-6}$, $\mu = 7.2$.

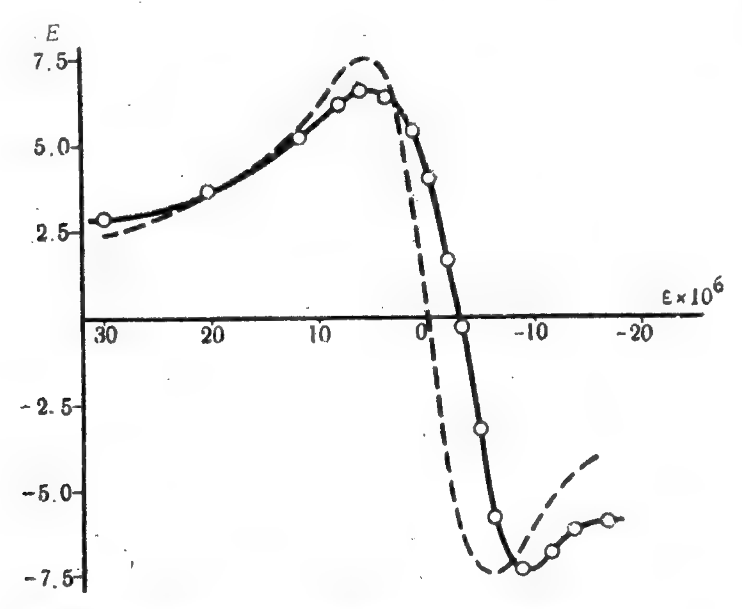
Пунктирная кривая—теоретическая, вычисленная по уравнению 15; кружки—опытные данные.

источник силы, тормозящей движение. В случае наличия адсорбированных молекул подобным же образом в растворе возникает разность концентраций. Непосредственно измерить разность потенциалов $\Delta\Phi$ между полюсами падающей капли не удается; однако легко показать, что ее существование приводит к появлению градиента потенциала E в столбе жидкости, в котором происходит падение капли, связанного с $\Delta\Phi$ простым соотношением

$$E = 2N\Delta\Phi \frac{s}{S}, \tag{16}$$

где N — число капель, которые одновременно приходятся на единицу длины трубки, s — сечение капли и S — сечение трубки, в которой происходит падение капель. Развитая нами теория позволяет выразить величину $\Delta\Phi$, а следовательно и E, через заряд капли, скорость падения, электропроводность и вязкость среды [7]. Как видно из фиг. 4, построенной по результатам измерений И. Багоцкой, выводы теории прекрасно подтвер-

ждаются на опыте. Интересно, что при возрастании заряда величина E сначала растет, а затем проходит через максимум и начинает падать. Это происходит вследствие того, что большие заряды тормозят движение поверхности и таким образом уничтожают причину возникновения разно-



Фиг. 4. Зависимость напряжения электрического поля E (в милливольтах на 1 см), возникающего при падении капель ртути в 0.03 N растворе KBr в глицерине, от ее заряда; $\varkappa = 1.1 \cdot 10^{-5}$, $\mu=9$.

Пунктирная кривая — теоретическая, кружки — опытные данные

сти потенциалов. Величина $\Delta\Phi$ оказывается неожиданно большой. Так, расчет и опыт показывают, что при падении капель ртути с радиусом 0.6 мм в 0.015 N растворе KBr в глицерине разность потенциалов между верхним и нижним концами капельки достигает 0.4 V.

4) *Измерение скорости химической реакции, протекающей на границе капли и окружающей среды.*

Этот метод, использующий несколько более сложные взаимоотношения, является, однако, наиболее плодотворным как в практическом, так и в теоретическом отношении. Поясним, на чем он основан. Предпо-

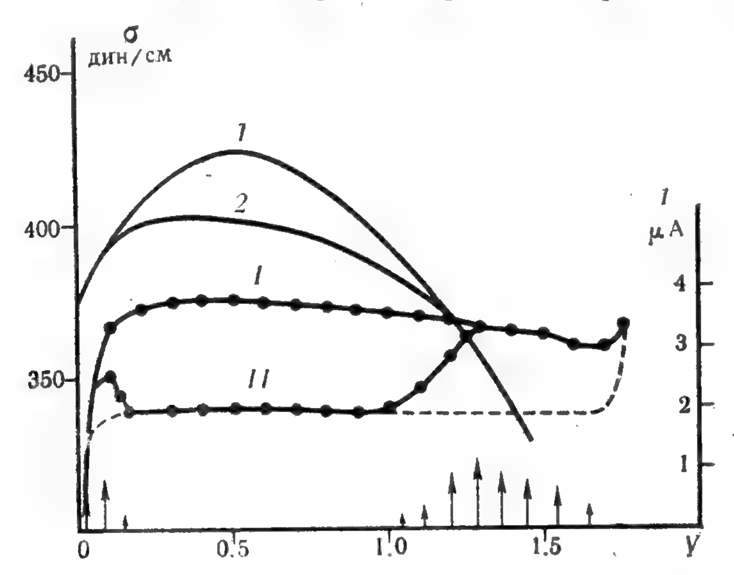
ложим, что на границе между каплей и окружающим раствором течет химическая реакция, скорость которой определяется диффузией какого-нибудь растворенного вещества к поверхности капли. Примером такой реакции может служить восстановление ионов ртути на катодно поляризованном ртутном капельном электроде. Тангенциальное движение поверхности, вызывающее размешивание раствора, увеличивает подачу вещества и ускоряет реакцию; в приведенном конкретном случае оно увеличивает, следовательно, силу тока, который может проходить через электрод. В случае вытекающей капли скорость подачи вещества в отсутствие тангенциального движения, как уже было указано выше, определяется уравнением Ильковича. Таким образом, наличие тангенциального движения приводит к появлению токов, превышающих токи, вычисленные по этому уравнению. В. Левичем было показано, что между скоростью тангенциального движения φ и вызванным этим движением током диффузии растворенного вещества существует соотношение

$$j = 0{,}85 \left(\frac{D\varphi}{a} \right)^{1/2} c, \qquad (17)$$

где j — число грамм-молей вещества, диффундирующих в единицу времени к единице поверхности, c — концентрация растворенного вещества, D — его коэффициент диффузии. Таким образом, измерение тока, текущего на капельный электрод, позволяет сделать количественные выводы о величине скорости движения поверхности. Явления адсорбции, приводящие к торможению движения, вызывают одновременно уменьшение силы тока вплоть до значений, соответствующих формуле Ильковича. Зависимость между силой тока, текущего на капельный электрод, и тангенциальным движением поверхности, связанным с вытеканием, была экспериментально обнаружена Крюковой и Кабановым [10] и изучена Крюковой [11] на большом опытном материале; теоретическое ее истолкование сделалось возможным благодаря изложенной здесь теории.

Использование капли в качестве электрода позволяет исследовать движение поверхности при различных потенциалах. В частности, таким способом удается определить тормозящее действие зарядов двойного слоя на движение поверхности, о котором мы уже говорили. Далее, варьируя разность потенциалов, можно изменять условия адсорбции органических молекул на границе ртуть | раствор. Мы здесь не можем оста-

навливаться на теории этого эффекта, хорошо известного из учения об электрокапиллярных явлениях. Укажем только, что адсорбция органических молекул происходит на границе ртуть | раствор лишь при неслишком больших значениях заряда поверхности в пределах определен-



Фиг. 5. Влияние адсорбции на зависимость тока от напряжения при наличии тангенциальных движений поверхности.

Кривая *I*—зависимость тока *I* от напряжения *V* для 1*N* $Na_2SO_4 + 3 \cdot 10^{-4}$ *N* $Hg_2^{\cdot\cdot}$; кривая *II*—то же в присутствии 0.1 мол. бутилового спирта; пунктирная кривая—то же после полного подавления движения добавкой желатины. Кривая *1*—зависимость пограничного натяжения σ от потенциала в 1*N* Na_2SO_4; кривая *2*—то же в присутствии 0.1 мол. бутилового спирта.

Стрелки показывают скорость движения поверхности ртути

ного интервала потенциалов, который всегда включает точку, соответствующую нулевому заряду поверхности, но несколько меняется от вещества к веществу. Это своеобразное сочетание гидродинамических и электрохимических факторов приводит к картине явлений, изображенной на фиг. 5. Кривая *I* выражает зависимость тока от наложенного напряжения в 1*N* растворе Na_2SO_4, содержащем небольшое количество ионов ртути, при наличии тангенциального движения поверхности. Пунк-

тирная кривая дает ту же зависимость для случая, когда тангенциальное движение полностью подавлено и капля растет только радиально (этого можно достигнуть, например, добавляя в раствор желатину). Кривая *II* получается в том случае, если к раствору Na_2SO_4 добавить бутилового спирта. Как видно из приведенных в верхней части чертежа кривых, выражающих зависимость пограничного натяжения ртуть | раствор от наложенной разности потенциалов для случая чистого раствора Na_2SO_4 и Na_2SO_4 + бутиловый спирт, последний адсорбируется только при поляризациях в пределах от — 0.2 до — 1.2 V. В этих же пределах исчезает тангенциальное движение поверхности, скорость которого изображена стрелками в нижней части чертежа, и кривая *II* сливается с пунктирной кривой тока на радиально растущую каплю. Этим методом можно детально изучить закономерности адсорбции органических молекул и вызываемого ею торможения движения.

Из уравнения (9) следует, что уменьшение скорости тангенциального движения определяется величиной коэффициента, который для случая торможения, вызываемого адсорбцией молекул из раствора, согласно уравнениям (7) и (10), равен

$$\gamma = \frac{2RT\Gamma_0^2\delta}{Dac}. \tag{18}$$

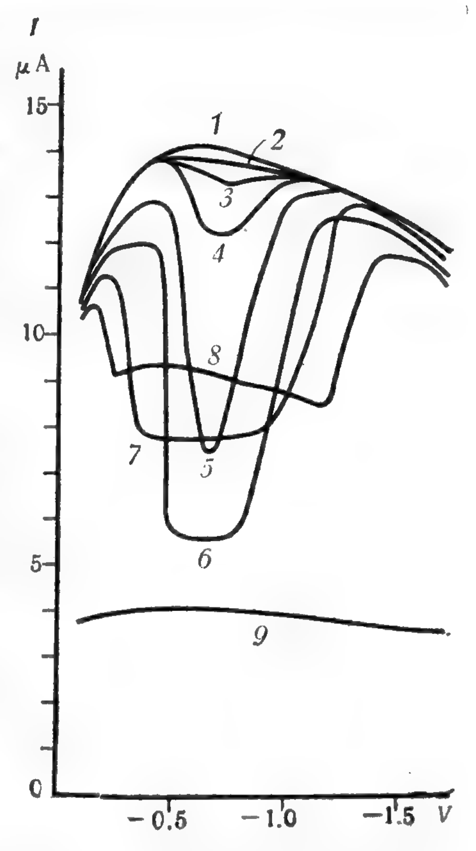
В выражение для γ входит квадрат величины адсорбции Γ_0. Как известно, при увеличении молекулярного веса в гомологических рядах адсорбируемость растет согласно правилу Траубе, увеличиваясь приблизительно в три раза при удлинении цепи на одну группу CH_2. Из уравнения (18) следует, что торможение должно расти в гомологическом ряду с увеличением длины цепи не в 3, а в 3^2, т. е. в 9 раз; иными словами, тот же эффект торможения должен получаться для вещества с более длинной цепью при девятикратном разбавлении. В табл. 1 приведены полученные Крю-

Таблица 1

Спирты	c	c_n / c_{n+1}
Пропиловый C_3H_7OH	$4\cdot10^{-2}$	
Бутиловый C_4H_9OH	$5\cdot10^{-3}$	8
Амиловый $C_5H_{11}OH$	$5\cdot10^{-4}$	10
Гептиловый $C_7H_{15}OH$	$7\cdot10^{-6}$	$(8.4)^2$

ковой значения концентраций нормальных спиртов, которые дают одинаковый эффект торможения, будучи добавлены к раствору $3\,N\,KCl + 3\cdot 10^{-4} N\,HgCl_2$, а именно снижают величину избыточного тока на одну треть. В последнем столбце даны отношения концентраций c для двух последовательных членов гомологического ряда. Как видно, отношение это действительно равно приблизительно 9.

Нужно отметить, что до сих пор не было известно эффектов, которые возрастали бы с такой быстротой при увеличении длины цепи; обнаружение такого эффекта — убедительное подтверждение правильности теории. Зависимость величины γ от концентрации определяется, согласно уравнению (18), изменением величины Γ_0^2/c. При малых значениях c величина Γ_0 приблизительно пропорциональна c, и поэтому γ растет с ростом c. При больших концентрациях Γ_0 достигает некоторого предела вследствие заполнения поверхности адсорбированными молекулами и дальнейшее увеличение c приводит уже к уменьшению γ. Таким образом, величина γ при увеличении концентрации должна сначала расти, достигнуть максимума и затем начать убывать. Этот второй интересный вывод из теории также нашел свое подтверждение на опыте, как видно из фиг. 6, на которой приведены полученные Крюковой кривые зави-



Фиг. 6. Зависимость тока I от напряжения V при наличии тангенциальных движений поверхности в растворах $3\,N\,KCl + 3\cdot 10^{-4} N\,HgCl_2$ в присутствии бутилового спирта.

Концентрация спирта: *1*—0; *2*—$2\cdot 10^{-3}$ мол.; *3*—$3\cdot 10^{-3}$ мол.; *4*—$6\cdot 10^{-3}$ мол.; *5*—$1.1\cdot 10^{-2}$ мол.; *6*—$3\cdot 10^{-2}$ мол.; *7*—10^{-1} мол.; *8*—$3\cdot 10^{-1}$ мол.; *9*—величина тока при полном торможении тангенциальных движений

симости силы тока от поляризации для растворов $3N$ KCl $+ 3 \cdot 10^{-4} N$ $HgCl_2$, содержащих добавки норм. C_4H_9OH в различной концентрации. Как видно из рисунка, наибольшее торможение соответствует некоторой промежуточной концентрации спирта, равной $3 \cdot 10^{-2}$ мол.; при дальнейшем увеличении концентрации область торможения продолжает расширяться, но самый эффект торможения уменьшается.

При выводе значений e и γ для торможения движения адсорбированными молекулами мы предполагали, что скорость выравнивания изменений поверхностной концентрации адсорбированных молекул определяется диффузией в объеме раствора. Однако такое предположение правильно только в том случае, если сами процессы адсорбции и десорбции протекают достаточно быстро. Нами были получены выражения для e и γ и в предположении, что медленной стадией является не диффузия к поверхности, а сам процесс адсорбции [4]. Эти выражения приводят к иным значениям для величины торможения и ее зависимости от концентрации, которые, однако, по крайней мере для случая спиртов нормального ряда, в отличие от уравнения (18), не согласуются с опытными данными. Таким образом, исследование этих эффектов позволяет решить, какая стадия адсорбционного процесса самая медленная, и дает новый способ изучения кинетики адсорбционных процессов в растворах.

Приведем еще один пример применения этих методов к кинетическим проблемам. Оказывается, что на основании такого рода измерений можно утверждать, что равновесное состояние двойного электрического слоя на границе ртуть | водный раствор электролита устанавливается за промежуток времени меньше 10^{-5} сек. Этот предел по крайней мере в 10 раз ниже того, который до сих пор удавалось установить с помощью наиболее совершенных непосредственных электрических измерений.[1] Сильная зависимость величины торможения органическими молекулами от длины цепи и в более общем случае от их способности адсорбироваться приводит к тому, что скорость движения поверхности оказывается чрезвычайно чувствительной к присутствию ничтожных количеств загрязнений.

[1] За такой же промежуток времени должно устанавливаться равновесное значение поверхностного натяжения чистой поверхности воды, что заставляет усомниться в правильности экспериментальных работ, в которых авторы пытались обнаружить различие между динамическим и статическим натяжением воды.

Поэтому для получения воспроизводимых значений при этих измерениях приходится подвергать исследуемые вещества гораздо более тщательной очистке, чем это необходимо в случае обычных адсорбционных измерений.

Соединения, хорошо адсорбирующиеся на границе ртуть | раствор, могут быть обнаружены по изменению формы кривой ток — напряжение на капельном электроде уже в ничтожно малых концентрациях. Например, β-нафтол оказывает уже заметное действие при концентрации 10^{-8} моля на литр. Кратковременное соприкосновение раствора с органическими веществами, например резиновой пробкой или бумажным фильтром, делает его уже совершенно непригодным для использования в этих измерениях, так как попадающих при таком соприкосновении в раствор ничтожных количеств загрязнений уже вполне достаточно, чтобы вызвать сильное торможение движения поверхности [11]. Применение этих методов позволяет также обнаружить изменения, которые происходят в растворах соединений, обычно считавшихся совершенно устойчивыми. Например, оказывается (по неопубликованным данным Т. А. Крюковой), что в растворах нормальных спиртов в воде медленно протекает какой-то химический процесс, который приводит к значительному увеличению их активности.

Несомненно, что эти методы в дальнейшем найдут себе применение не только для исследования механизма адсорбционных явлений, но и для аналитических целей.

Полученные результаты проливают свет на ряд смежных вопросов физической и коллоидной химии; из них назовем устойчивость пен, многие явления при полярографическом анализе, поведение некоторых коллоидных систем.

В настоящей статье я не коснулся вопроса о движении на границах раздела жидких фаз под влиянием электрического поля, которому также было посвящено несколько наших работ [12]; такое движение характеризуется рядом интересных особенностей.

Мне кажется, что дальнейшее развитие работ в этих направлениях приведет к созданию новой главы физической химии — физико-химической гидродинамики, основанной в очень значительной мере на теоретических и экспериментальных результатах, полученных советскими учеными.

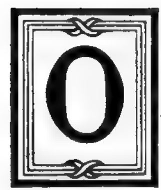
ЛИТЕРАТУРА

1. *Р. Н. Иванов.* Изв. АН СССР, серия геогр., 1937; стр. 325, 1938, стр. 29; *В. В. Шулейкин.* Там же, 1937, стр. 345.
2. *В. Г. Левич.* К теории поверхностных явлений. М., 1941.
3. *W. Gibbs.* Equilibrium of heterogeneous substances, стр. 301.
4. *А. Н. Фрумкин* и *В. Г. Левич.* ЖФХ (в печати); *В. Г. Левич.* ЖФХ, т. 21, 1947, стр. 689.
5. *Д. Л. Талмуд* и *С. Е. Бреслер.* Поверхностные явления. Л.— М., 1934.
6. *В. Г. Левич.* ЖФХ, т. 18, 1944, стр. 335.
7. *А. Н. Фрумкин* и *В. Г. Левич.* ЖФХ, т. 21, 1947, стр. 953.
8. *А. Н. Фрумкин.* Добавление к переводу книги: Кольтгоф и Лингейн. Полярография, М., 1947.
9. *И. А. Багоцкая* и *А. Н. Фрумкин.* ДАН СССР, т. 55, 1947, стр. 135.
10. *Т. А. Крюкова* и *Б. Н. Кабанов.* ЖФХ, т. 15, 1941, стр. 475.
11. *Т. А. Крюкова,* ЖФХ, т. 20, 1946, стр. 1179; т. 21, 1947, стр. 365.
12. *А. Н. Фрумкин* и *В. Г. Левич.* ЖФХ, т. 19, 1945, стр. 573; *А. Н. Фрумкин,* Изв. АН СССР, серия хим., 1945, стр. 223.

Академик

П. А. Ребиндер

ФИЗИКО-ХИМИЧЕСКИЕ ИССЛЕДОВАНИЯ ПРОЦЕССОВ ДЕФОРМАЦИИ ТВЕРДЫХ ТЕЛ



дна из основных задач современной физической химии состоит в установлении связи между физическими свойствами тел и их химическим составом и строением. Решение этой задачи позволяет управлять свойствами тел, например, изготовлять металлические сплавы с заранее заданными свойствами соответственно их прикладному назначению и проектировать материалы для современной техники как некое сложное и тонкое сооружение.

Основными свойствами твердых тел — отдельных крупных кристаллов и кристаллических агрегатов с зернами различных размеров являются, несомненно, их механические свойства. Все могущество современной техники — величественные здания и сооружения, точные и мощные машины — основано на использовании механических свойств твердых материалов (металлов, камня, бетона, дерева, пластических масс, резины и др.), из которых они построены. Механические свойства тел — упругость, пластичность, вязкость, прочность—и закономерности, определяющие эти свойства в различных условиях, представляют также выдающийся интерес для современной науки. Эти свойства наиболее непосредственно определяются молекулярными силами, силами сцепления между частицами тела и его строением из этих частиц — молекул, атомов или ионов. Исследование механических свойств твердых тел и процессов дефор-

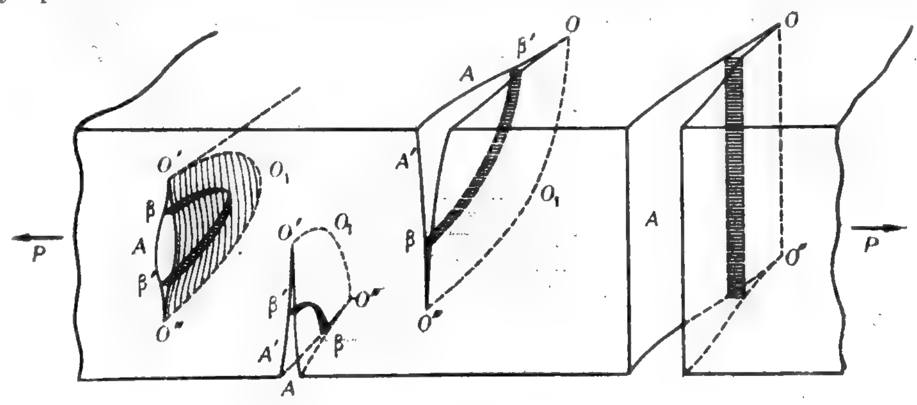
мации в этих телах приводит к существенным выводам об особенностях их строения и позволяет заранее судить о способности твердых материалов поддаваться механической обработке различного вида (резание металлов или обработка их давлением — прокатка, волочение, штамповка, разрушение горных пород при бурении и взрывании, измельчение минерального сырья или зерна при помоле).

В механике и физике, а тем более в различных разделах механической технологии, механические свойства твердых тел изучались без должного внимания к физико-химическим факторам. Во всяком случае всегда считали, что для твердых тел обычных размеров механические свойства при данной температуре и внешнем давлении определяются только внутренней природой самого тела. Как мы увидим далее, систематическое исследование влияния состава внешней среды на процессы деформации твердых тел не только привело к установлению новых интересных явлений и закономерностей, но и позволило по-новому рассматривать самый механизм процессов деформации кристаллов в наиболее общих случаях.

При этом решающее значение оказывает содержание в окружающей среде поверхностноактивных веществ. Молекулы этих веществ адсорбируются поверхностями деформируемого твердого тела, т. е. прилипают к этим поверхностям под действием свободных молекулярных сил, тех самых сил, которые вызывают явления смачивания, паяния и сварки металлов.

Работами, ведущимися в нашей лаборатории уже около двадцати лет [1], установлено весьма общее явление понижения сопротивления твердых тел упругим и пластическим деформациям, а также разрушению, под влиянием адсорбции поверхностноактивных веществ из окружающей среды. Эти явления адсорбционного облегчения деформаций или адсорбционного понижения прочности (твердости) образуют в настоящее время новую область физико-химической механики твердых тел. Еще раньше мы считали, что эти явления обусловлены облегчением возникновения и развития микрощелей в поверхностных слоях деформируемого тела. В дальнейшем удалось в значительной степени выяснить механизм этих удивительных явлений. Он в основном состоит в облегчении развития микрощелей, которые возникают на основе слабых мест — дефектов в кристаллической решетке, играющих роль зародышей [2]. При этом чувствительны к адсорбции только те микрощели, устья которых выходят на внешнюю поверхность кристалла, а «тупиковые» части остаются внутри тела. Такие

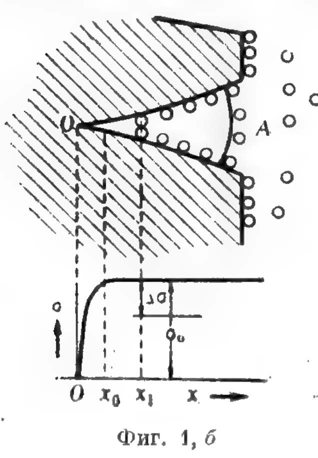
щели всегда имеют клиновидное сечение (фиг. 1, *а* и *б*) и возникают при всяком объемно-напряженном состоянии тела, когда на данную элементарную площадку в любой точке тела действуют, кроме нормальных (растягивающих), еще и тангенциальные напряжения. Естественно, что развитию таких щелей благоприятствуют растяжения и, наоборот, препятствуют сжатия: при всестороннем сжатии развитие микрощелей вообще устраняется.



Фиг. 1, *а*

Адсорбция (Γ) вызывает понижение поверхностного натяжения σ на поверхностях твердого тела на величину $\sigma_0 - \sigma_\Gamma$, причем $\sigma_0 - \sigma_\Gamma = RT\Gamma$ в предельной области весьма малых Γ. Действие адсорбционных слоев сводится к их двухмерной миграции по поверхностям в устьях микрощелей до стерического препятствия, обусловленного собственными размерами адсорбирующихся молекул и соответствующего критической толщине зазора в микрощели. В этом критическом зазоре естественная граница адсорбционного слоя ($\beta\beta'$ на фиг. 1, *а*) образует линейный барьер, на каждую единицу длины которого действует двухмерное давление $\sigma_0 - \sigma_\Gamma$ в сторону его дальнейшего продвижения в глубь микрощели, способствуя таким образом развитию микрощелей (при постоянстве внешних усилий) и нарастанию деформаций. Эффект этого давления эквивалентен увеличению внешнего усилия F на величину $\Delta F \sim \sigma_0 - \sigma_\Gamma$, заменяющую собой действие адсорбционных слоев и как бы являющуюся их механическим эквивалентом.

При адсорбции из смачивающей жидкой среды (если деформируемое тело помещено в раствор поверхностноактивного вещества) жидкость проникает в устья микрощелей под влиянием капиллярного давления. Однако при этом с мениска отрываются молекулы наиболее поверхностноактивного компонента, которые мигрируют впереди и покрывают поверхности щели со значительно большей скоростью, чем всасывание жидкости в целом, испытывающей вязкое сопротивление. В части микрощелей, заполненной жидкостью (вблизи устья), тонкая пленка жидкости, заполняющая зазор щели, может производить дополнительное расклинивающее давление. Это давление служит мерой лиофильности твердого тела, его сродства к данной жидкости, а потому может усиливаться при адсорбции в результате соответствующей ориентации адсорбционного слоя. Наличие такого расклинивающего давления было непосредственно установлено в опытах Б. В. Дерягина и его сотрудников, например, между плоскопараллельными пластинками с изменяющимся зазором, помещенными в данную жидкость [3].



Фиг. 1, *б*

Однако такое расклинивающее давление в устьях обычно, повидимому, значительно меньше двухмерного давления у передней границы адсорбционного слоя. Значение последнего становится особенно ясным в установленных нами сильных эффектах адсорбционного облегчения деформаций слюды [4], деформируемой изгибом или растяжением в воде, при добавлении к воде спиртов, адсорбирующихся слюдой из водного раствора, как это было показано прямыми измерениями. При этом молекулы спирта адсорбируются на поверхности слюды полярными группами к слюде, а углеводородными цепями — в водный раствор, что уменьшает смачивание слюды водой и вместе с тем снижает расклинивающее давление; вместе с тем добавки спирта значительно понижают поверхностное натяжение, адсорбируясь на границе слюда — водный раствор и создавая высокое двухмерное давление.

Каждой величине среднего напряжения или, что то же самое, в упругой области, каждой величине деформации соответствует свой равновес-

ный размер микрощелей или, строго говоря, своя кривая распределения микрощелей с определенным вероятнейшим их размером.

Уже указывалось [5], что, кроме облегчения развития каждой отдельной микрощели под влиянием проникновения в нее адсорбционных слоев, эффект адсорбционного облегчения деформаций обусловлен также и увеличением числа микрощелей, развивающихся на единицу объема деформируемого тела. Постепенное во времени смыкание микрощелей от «линии тупиков» $O' O_1 O''$ к устью AA' (фиг. 1, *а*), под влиянием сил сцепления в твердом теле, аналогично коалесценции — слиянию двух капель жидкости (например, ртути) под действием поверхностного натяжения. Вблизи тупиков микрощели имеют зародышевую часть, в которой силы сцепления нарастают до нормальной наибольшей величины, соответствующей обращению зазора в ноль, т. е. переходу к нормальной структуре тела в самом тупике. Соответственно этому поверхностная энергия (фиг. 1, *б*) быстро нарастает с увеличением ширины зазора в щели от тупика к устью по направлению оси x до нормального наибольшего значения, соответствующего свободной внешней поверхности тела (σ_0).

Еще в опыте И. В. Обреимова с расщеплением слюды по спайности в вакууме [6] было показано, что расщеп при вдвигании стеклянного клина вновь обратимо смыкается при выдвигании его, и последующее расщепление требует затраты почти той же механической работы. На этом основан предложенный Обреимовым остроумный метод измерения поверхностного натяжения грани спайности слюды.

Расщепление слюды в подобных опытах вполне аналогично раздвиганию клиновидных микрощелей, развивающихся, по нашим представлениям, и в области упругих деформаций без всякого разрушения твердого тела. Как было любезно указано Я. И. Френкелем, такие представления об обратимо смыкающихся микрощелях с зародышевой, тупиковой, частью приводят к устранению важных недостатков в известной теории Гриффитса и открывают пути к существенному ее усовершенствованию.

Смыкание микрощелей после разгрузки должно происходить и при пластических деформациях с той разницей, что после скольжения-сдвига смыкание происходит по новым местам. Достижение предела прочности, т. е. разрыв, соответствует полному развитию одной из микрощелей, пересекающей все поперечное сечение образца. При этом происходит мгно-

венная разгрузка частей тела по обе стороны разрыва, все остальные микрощели смыкаются, и их поверхностная энергия, как и вся упругая энергия тела, после разрыва рассеивается в виде тепла.

Если бы развивающиеся микрощели были с плоскопараллельным зазором, то адсорбция из внешней среды частиц, обычно по размеру бóльших, чем структурные элементы самого кристалла, могла бы происходить лишь после того, как сцепление преодолено по всей щели и разрыв уже произошел. В клиновидных же, т. е. реальных, микрощелях, устья которых выходят на внешнюю, свободную поверхность твердого тела, адсорбирующиеся молекулы свободно проникают в эти устья и, достигнув такого места в клиновидной щели, где собственные их размеры уже препятствуют дальнейшему проникновению, образуют таким образом «границу адсорбционного слоя». Вдоль этой границы (ββ' на фиг. 1, *а*) адсорбирующиеся молекулы должны оказывать давление, стремящееся далее раздвинуть трещину и по своей величине пропорциональное понижению поверхностного натяжения. Этот вывод термодинамически необходим, а уже описанное действие в микрощели вполне аналогично действию двухмерного давления мономолекулярного слоя органических поверхностно-активных веществ на поверхности чистой воды, толкающих барьер, ограничивающий этот слой в известных опытах Лангмюра. У нас роль барьера играет критическая ширина зазора в клиновидной щели, приблизительно равная удвоенному диаметру адсорбирующихся молекул (см. фиг. 1, *б*). Таким образом, как уже было указано, проникновение адсорбирующихся молекул в устье микрощелей и создает дополнительные раздвигающие усилия, что эквивалентно некоторому увеличению внешних сил; при заданных же механических условиях (т. е. заданных внешних силах) этот эффект вызывает возрастание величины деформации в упругих телах или скорости деформации — в телах упруго-пластических.

Возможные случаи взаимодействия твердого тела с адсорбирующимися веществами из внешней среды разделяются на две группы [2], если пользоваться следующей простейшей схемой. Пусть C — работа сил сцепления в кристалле, энергия связи частиц твердого тела; A — работа адсорбционных сил, энергия связи адсорбирующихся молекул с поверхностью тела; E — работа внешних сил при деформации тела в расчете на данную микрощель. Тогда возможны два случая:

$$1)\ A > C, \qquad 2)\ A < C.$$

В первом случае силы адсорбции превышают силы сцепления в твердом теле хотя бы на некоторых участках, что соответствует явлениям набухания и коллоидного растворения — пептизации сравнительно мягких и не однородных по структуре и составу тел. Для нас более интересен второй случай, так сказать, истинно твердых тел, для которых силы сцепления в теле заведомо превышают силы адсорбции. В этом случае проникновения адсорбирующихся молекул в недеформированное тело не будет, что и было нами показано для кристаллов кальцита, слюды или металлов. Однако наши опыты показали, что проникновение даже крупных молекул поверхностноактивных органических веществ может происходить в устья микрощелей при деформировании, когда работа деформации E вместе с работой адсорбционных сил A превышает энергию сцепления: $A + E > C$, т. е. при условии, что $E > C - A$. Отсюда следует, что при разгрузке тела вновь $A < C$, что соответствует самопроизвольному смыканию микрощелей, начиная с тупика — зародышевого конца щели, с вытеснением адсорбционных слоев, подобно тому как передвиганием барьера Лангмюра по поверхности воды можно очищать ее от мономолекулярного слоя. В нашем случае роль работы внешних сил, приложенных к барьеру, играет работа сил сцепления в зародышевой части микрощелей. Проникшие в микрощели в процессах деформации адсорбционные слои замедляют, таким образом, смыкание микрощелей или, что то же самое, постепенное исчезновение упругой деформации после разгрузки тела.

Интересно отметить, что при контакте двух частиц в суспензии или коллоидном растворе, коагулировавших в результате действия сил сцепления, всегда образуется клиновидный зазор вдоль периметра контакта. Действие пептизаторов, адсорбционные слои которых вызывают распад коагуляционных агрегатов на свободные первичные частицы, можно свести к указанному раздвиганию клиновидной микрощели. Так как при этом коагуляционное сцепление обычно бывает слабым, явление пептизации относится к первой группе по только что рассмотренной схеме.

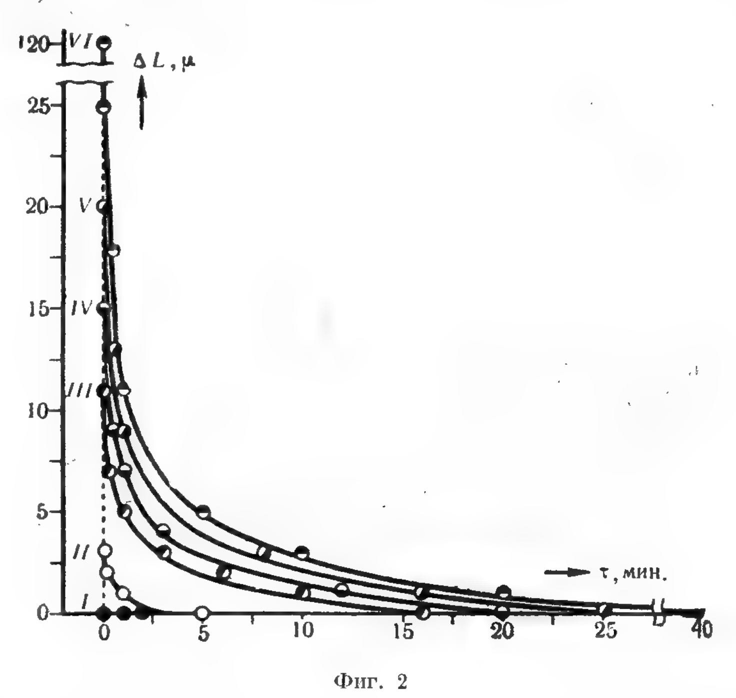
Эти общие представления были подтверждены нашими работами. Здесь мы рассмотрим опыты над упругими деформациями листков слюды и упруго-пластическими деформациями гипса, а также над пластическим течением монокристаллов олова и свинца.

Процессы адсорбционного впитывания поверхностноактивных веществ в микрощели при деформации и их вытеснения из микрощелей при смыкании после разгрузки протекают во времени с определенной скоростью и поэтому должны вызывать явления особого ярко выраженного упругого последействия. Эти явления и были действительно установлены в нашей лаборатории Г. И. Логгиновым в ряде простых опытов изгиба и растяжения кристаллов слюды.

Работами ряда исследователей и прежде всего классической работой А. Ф. Иоффе [7] было показано, что однородные кристаллы не обнаруживают упругого последействия. И действительно, в наших опытах для хороших образцов слюды в вакууме, как и в сухом воздухе, упругого последействия при нагрузке или разгрузке практически не было. Однако при достаточно длительном предварительном действии деформирующей силы небольшое последействие разгрузки появилось даже и в воздухе, что могло быть вызвано остаточными следами влажности. При деформировании же слюды в парах воды, или, что то же самое, в жидкой воде, всегда возникает резко выраженное длительное упругое последействие, сравнительно кратковременное после разгрузки и весьма длительное при нагружении. Это последействие не обычное: оно не оставляет, как обычно в реальных неоднородных упругих телах, небольших долей от величины мгновенной деформации, т. е. деформации, возникающей в теле со скоростью звука. «Адсорбционное» последействие нагрузки в слюде (см. фиг. 3) может выражаться в постепенном нарастании деформации, соответствующем увеличению ее в два раза и более по сравнению с начальным, «мгновенным», значением.

На фиг. 2 показано упругое последействие разгрузки листочков слюды, растянутых постоянным усилием (2.5 кГ) в присутствии адсорбирующегося вещества (воды) после мгновенного снятия нагрузки в момент времени $\tau = 0$. По оси оридинат отложены абсолютные удлинения ΔL в микронах, по оси абсцисс — время последействия τ в минутах. Кривая *I* (черные точки, совпадающие с осью абсцисс) иллюстрирует полное отсутствие последействия в предельно сухом воздухе, практически мгновенное исчезновение всей деформации, равной 120 μ после разгрузки. Кривые *II*, *III*, *IV*, *V*, *VI* иллюстрируют кинетику упругого последействия разгрузки при растяжении такой же полоски слюды в воздухе, насыщенном парами воды при 20° С, или при смачивании водою

для следующей продолжительности пребывания в адсорбционно-активной среде под той же нагрузкой (2.5 кГ): *II* — 2 сек., *III*—1 мин., *IV*—5 мин., *V*—10 мин., *VI*—20 мин.

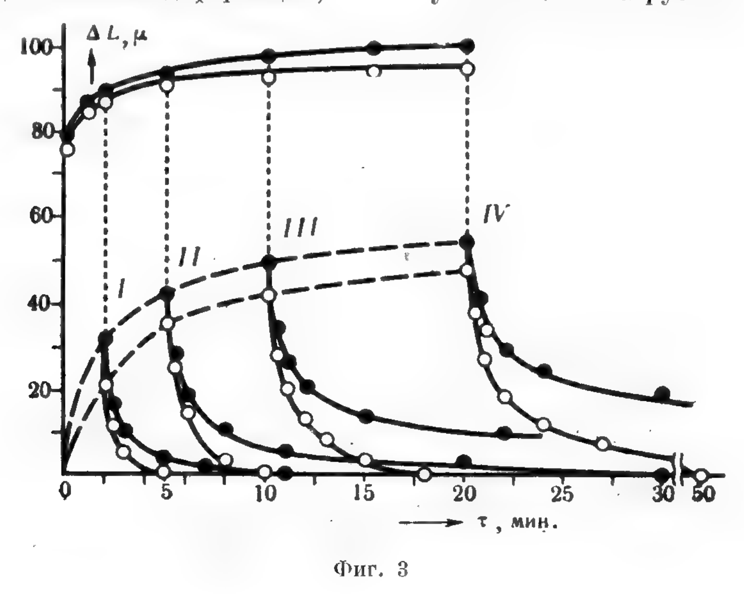


Фиг. 2

На фиг. 3 представлено упругое последействие нагрузки и разгрузки при растяжении полосок слюды в воде (белые кружки) и в водном растворе изоамилового спирта (черные кружки) при концентрации последнего 0.3 моль/л (близкой к насыщению). Нагрузка — практически мгновенная (как и разгрузка), во всех случаях одинаковая (2.5 кГ). По оси ординат — абсолютные деформации удлинения ΔL в микронах, по оси абсцисс — время последействия τ в минутах. Кривые последействия разгрузки измерялись на различных стадиях последействия нагрузки (*I*—2 мин., *II*—5 мин., *III*—10 мин., *IV*—20 мин.).

Последействие еще значительно усиливается при добавлении к воде, в которую погружается деформируемый кристалл, веществ, адсорбирующихся слюдой из водного раствора, например, спиртов — октилового

или изоамилового. Время, потребное для достижения предельной деформации изгиба, доходит в этих случаях до трех суток. При этом конечный модуль упругости слюды в результате адсорбции значительно понижается — в два раза и более. Эффекты становятся более резко выраженными с возрастанием величины деформации, т. е. с увеличением нагрузки или вре-



Фиг. 3

мени последействия. Кривые последействия довольно хорошо следуют простейшему логарифмическому закону

$$\Delta L = a + b \log \tau,$$

вытекающему из классической теории Больцмана. Можно показать, что адсорбционные слои, проникающие в микрощели деформируемого тела, вызывают его «адсорбционное размягчение» не только активным раздвиганием микрощелей, но и, так сказать, пассивно, замедляя их смыкание. Это обнаруживается при периодически налагаемых напряжениях с достаточно высокой частотой, чтобы обратное смыкание микрощелей за разгрузочный период не успевало произойти.

Нами было показано, что адсорбционные эффекты особенно резко и непрерывно нарастают, если при растяжении листка слюды постоянной

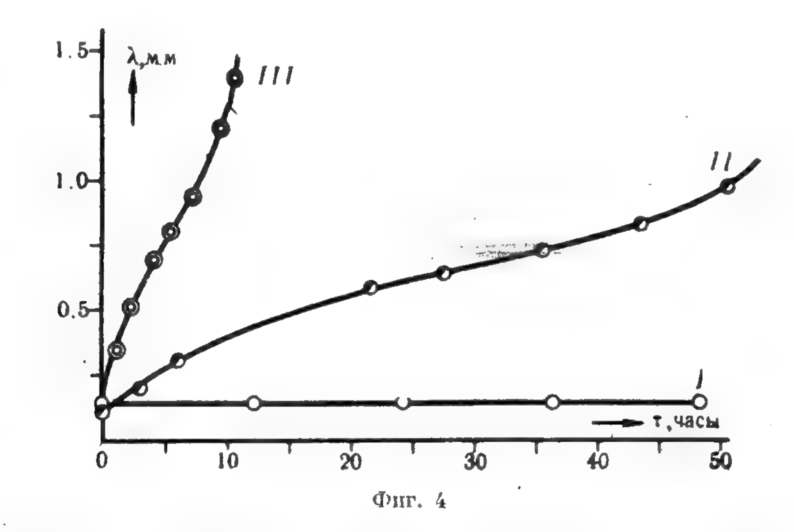
силой налагать на него еще дополнительное периодическое усилие при помощи простого электромагнитного приспособления. При этом легко оценить величину упругого гистерезиса [2] по расхождению кривых нагружения и разгружения (с постоянной заданной скоростью прироста или снятия нагрузки). Площадь петли гистерезиса, образованная этими кривыми и измеряющая рассеянную энергию за цикл, непрерывно нарастает при таком расшатывании кристалла и значительно увеличивается при действии адсорбирующихся веществ (например, изоамилового спирта, добавляемого к воде, в которой проводится весь опыт). В сухом воздухе в этих условиях упругий гистерезис практически отсутствует.

Особенно интересно отметить, что все эти явления наблюдаются в области упругих деформаций: по окончании каждой серии исследований, снятия нагрузки и при достаточно длительной выдержке деформации кристалла исчезают полностью.

Исходя из изложенной выше картины развития микрощелей в области объемно-напряженного состояния, можно было предполагать, что при растяжении таких же листков слюды, но с защитными краями, значительно выступающими за края зажимов, адсорбционных эффектов не будет или они будут сильно ослаблены. Это и было подтверждено в наших опытах, подтвердивших также схему развития поперечных микрощелей на фиг. 1, *а*.

Таким образом, под влиянием незначительных изменений состава окружающей среды исследуемое твердое тело не только легче деформируется, но и переходит при деформации в другую категорию: из практически истинно упругого оно становится телом с явственно выраженным упругим последействием, а в других случаях и с более заметной релаксацией, т. е. с конечным периодом релаксации. Это означает, что при проникновении адсорбирующихся веществ в деформируемые твердые тела последние по своим механическим свойствам все более приближаются к жидкостям. Для тел, самопроизвольно набухающих уже в недеформированном состоянии (первая группа тел в указанной нами схеме), подобное действие хорошо известно и носит название пластифицирования или мягчения. Теперь мы видим, что подобного же рода процессы наблюдаются и для истинно твердых тел (вторая группа), но только с участием механической работы деформации.

Рассмотренная картина делает понятным тот факт, что адсорбционное действие среды сказывается в постепенном размягчении образцов твердого тела любых размеров. Требуется лишь, чтобы проникновение адсорбционных слоев в деформируемый объем успевало произойти за время деформирования. Как известно, впитывание воды в кристаллы обычной каменной соли было доказано оптическими методами. Такое впитывание, по нашим представлениям, могло бы, очевидно, значительно возрасти при деформации кристалла даже в упругой области.

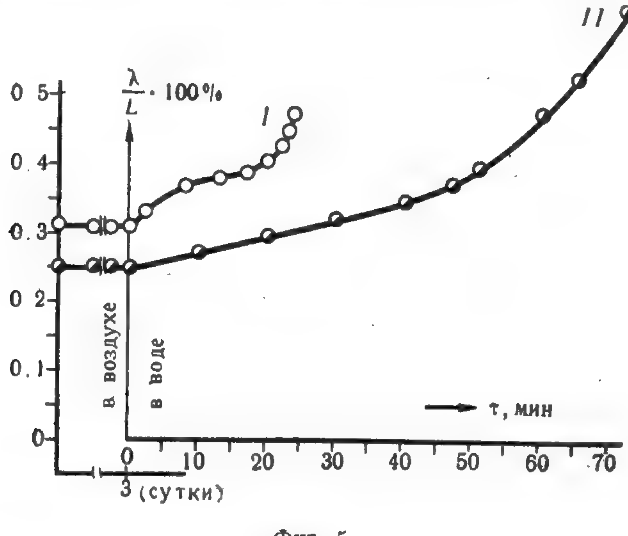


Фиг. 4

Изложенные представления были использованы нами для объяснения ползучести, возникающей у строительных гипсовых материалов под влиянием влаги. Такие материалы (обычный и высокопрочный демпферный гипс, гипсобетон) и строительные детали из них, обладая вполне удовлетворительными механическими свойствами в воздушно-сухом состоянии, под действием изгибающих нагрузок, лежащих значительно ниже предела прочности, обнаруживают нарастание деформаций с увеличивающейся скоростью («катастрофическую» ползучесть) под влиянием увлажнения. Это происходит, например, при воздействии насыщенного водяного пара, конденсирующегося в пористом гидрофильном материале, каким является гипс, или при смачивании водой, в частности под действием дождя. Замечательное явление своеобразной ползучести при увлаж-

нении наглядно проявляется в опытах с изгибом гипсовых брусочков (фиг. 4 и 5).

На фиг. 4 показано влияние смачивания водой и водными растворами электролитов на деформации изгиба гипсовых призматических брусков квадратного сечения. По оси ординат отложена стрела прогиба λ в миллиметрах, по оси абсцисс — время τ в часах пребывания под постоянной нагрузкой (максимальное нормальное напряжение $P_N = 16$ кГ/см²). Пря-



Фиг. 5

мая *I* иллюстрирует постоянство деформации при изгибании в воздухе (отсутствие ползучести); *II* — кривая ползучести при обильном смачивании образца водой; *III* — то же при смачивании конц. 10—30% водным раствором $CaCl_2$. Кривые заканчиваются наступлением разрыва.

На фиг. 5 показана картина возникновения ползучести под действием смачивания водой при изгибе брусков из обыкновенного поликристаллического гипса (кривая *I*) и полосок, вырезанных вдоль спайности из однокристального гипса (кривая *II*).[1] По оси ординат отложена приведенная стрела прогиба (в процентах от длины изгибаемого образца), по оси

[1] Бруски квадратного сечения — толщиною 15 мм и длиною 90 мм, полоски однокристального гипса — толщиною 1.0 — 1.5 мм, шириною 8 — 10 мм и длиною 40 мм.

абсцисс — время нарастания деформации после начала смачивания водой (в минутах) при изгибе постоянной нагрузкой: для поликристаллического гипса $P_N = 40$ кГ/см², для однокристального гипса $P_N = 100$ кГ/см².

До начала смачивания водой образцы под той же постоянной нагрузкой проверялись на отсутствие заметной ползучести в течение трех суток на воздухе. Кривые заканчиваются разрывом образца.

Наши опыты доказали несостоятельность объяснения ползучести гипса растворяющим действием воды. В действительности добавление к воде электролитов с одноименными по отношению к гипсу ионами хлорида кальция, сульфата магния (даже в высоких концентрациях) значительно повышает ползучесть образца, тогда как скорость растворения и соответственно растворимость гипса в таких водных растворах падают в десятки раз. Ясно, что добавляемые к воде электролиты, как и типичные поверхностноактивные вещества, действуют в этом случае адсорбционно, являясь своеобразными понизителями твердости [5]. Соответствующие опыты были проведены также и на однокристальных пластинках, вырезанных из расколотых по спайности крупных кристаллов гипса. При этом были установлены и исследованы подобные же явления ползучести при смачивании водой и водными растворами (фиг. 5), однако с той разницей, что в случае однокристального гипса (беспористого материала) скорость деформации убывает со временем, и явление становится более похожим на последействие, хотя при больших нагрузках оно вновь принимает характер типичной ползучести.

Пористость строительных гипсовых материалов, как это и наблюдается, должна значительно увеличить адсорбционные эффекты, так как благодаря пористости поверхностноактивные вещества (вода) уже при смачивании проникают в глубь недеформированного твердого тела в будущую зону деформаций.

На основании этих представлений о природе ползучести были разработаны пути ее предотвращения методом гидрофобизации гипса и введения крупнозернистых армирующих наполнителей (при затворении гипсового теста), препятствующих развитию пластических деформаций в материале. Эти исследования проводились в контакте с Институтом стройтехники Академии архитектуры.

За последнее время в разных областях возникают взгляды, близко соответствующие нашим представлениям. Так, в 1944 г. Орован [8] опуб-

ликовал заметку об усталости стекол под нагрузкой: при длительном нагружении или нарастании нагрузки с очень малой скоростью стекла обнаруживают значительно меньшую прочность, чем в обычных кратковременных испытаниях. Эти явления обусловлены, повидимому, адсорбцией воды из влажной атмосферы на поверхностях микрощелей, развивающихся в гидрофильном стекле. Подобные явления должны обнаруживаться и по нарастанию усталости у металлов, подвергаемых длительному периодическому деформированию, если такие опыты проводятся в присутствии адсорбирующихся веществ. Многие явления коррозионной усталости можно пытаться свести к такому механизму, особенно если учесть новые вторичные явления (см. ниже).

Изложенное относилось прежде всего к упругим деформациям. Картина, наблюдаемая при пластических деформациях, отличается лишь тем, что смыкание микрощелей в решетке происходит после сдвига, так сказать, по новым местам и, следовательно, имеет место не только после разгрузки, но также и в самом процессе деформации (например, при пластическом течении с постоянной скоростью), как это будет подробнее изложено нами в дальнейшем сообщении.

Наибольший интерес представляет исследование адсорбционного эффекта облегчения деформаций на отдельных кристаллах, а по отношению к пластическим деформациям — на металлических монокристаллах [9—10]. Такие исследования проводились в нашей лаборатории В. И. Лихтманом на монокристаллах олова, цинка, свинца и алюминия. На них удалось с особенной ясностью установить основные закономерности влияния поверхностноактивных веществ на механизм пластического течения.

Монокристаллы олова высокой чистоты получались по методу Капицы диаметром 0.8—1 мм. Для получения более однородной воспроизводимой поверхности они протравливались в слабой соляной кислоте. Для описываемых ниже опытов выбирались только монокристаллы, ориентация плоскостей скольжения которых была близка к оптимальной и соответствовала углам с осью кристалла около 45° [10]. Для правильного сопоставления, во избежание ошибок, монокристалл разрезался на 3—4 образца длиною около 30 мм каждый. Полные диаграммы растяжение — напряжение измерялись на приборе типа Поляни с постоянной скоростью относительного удлинения $V=\frac{d\varepsilon}{d\tau}=4.8\%$ мин.$^{-1}$ Растягивающее усилие

35*

измерялось по стреле прогиба жесткой стальной пластинки посредством двойного оптического рычага.

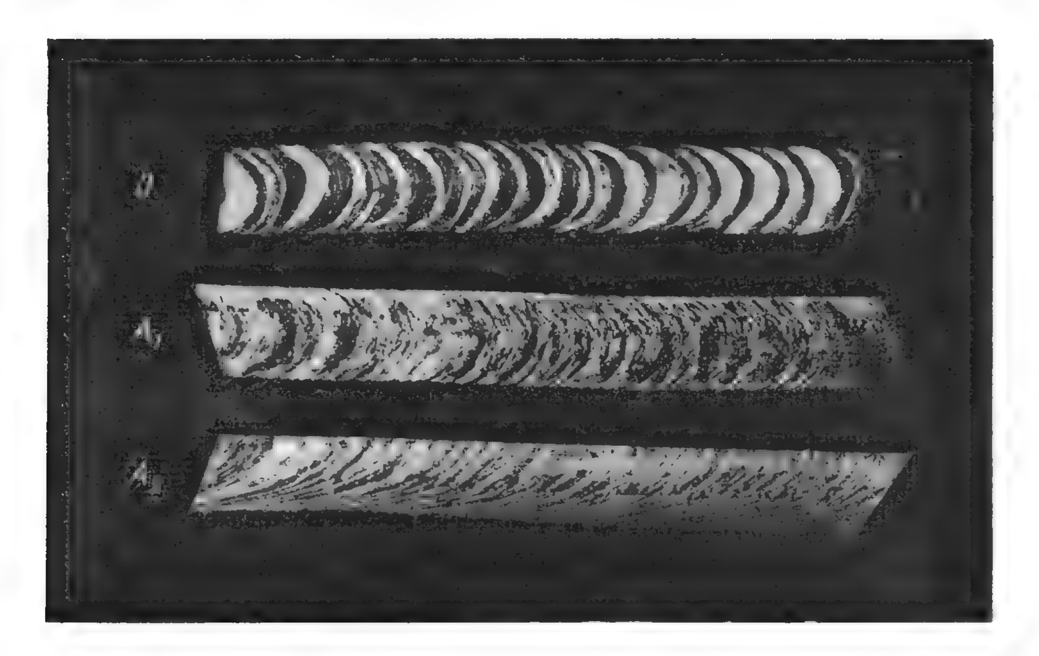
Относительное удлинение определялось по скорости вращения направляющего винта прибора, вращаемого от электромотора через редуктор, с поправкой на деформацию динамометрической пластинки, т. е. на смещение верхнего зажима.

Во избежание искажающих напряжений в зажимах образцы по 10 мм с каждого конца заплавлялись сплавом Вуда в стеклянные трубки с оттянутыми кончиками, которые и закреплялись в зажимах прибора. Предварительно устанавливалось, что диаграммы растяжения в воздухе и в неполярном вазелиновом масле, т. е. чисто углеводородной жидкости, тщательно очищенной от поверхностноактивных примесей, полностью совпадают.

Нами проводилось исследование влияния добавок очень чистой олеиновой кислоты к неполярному вазелиновому маслу в различных концентрациях на диаграмму растяжения вплоть до значительных удлинений [11] $\varepsilon_F = \frac{L - L_0}{L_0} 100 = 200 \div 300\%$.

Для изучения характерных поверхностных структур, возникающих при пластическом растяжении монокристаллов олова, нами совместно с Фотокинолабораторией АН СССР (С. Ю. Чураковский) был применен метод ускоренной микрокиносъемки (64 кадра в 1 сек.) в отраженном свете с помощью металломикроскопа; этот метод дал возможность зарегистрировать весь процесс возникновения пластического течения монокристалла как в инактивной, так и в активной среде. Такой метод исследования позволил установить, что уже на начальных стадиях деформации в присутствии поверхностноактивных веществ появляется очень тонкая структура пачек скольжения. Возникающие пачки скольжения имеют значительно меньшую толщину — порядка 1 μ (фиг. 6, A_1, A_2), чем при тех же скоростях растяжения в обычных условиях в инактивной среде, когда толщина пачек скольжения составляет в среднем 52 μ. В последнем случае в условиях значительного разброса этой величины очень тонкие пачки скольжения совсем отсутствуют, так что вся картина, обнаруживаемая микрокиносъемкой, вполне соответствует классической модели скошенной «стопки монет» Поляни (фиг. 6, *O*). Таким образом, действие поверхностноактивных веществ становится видимым в прямом

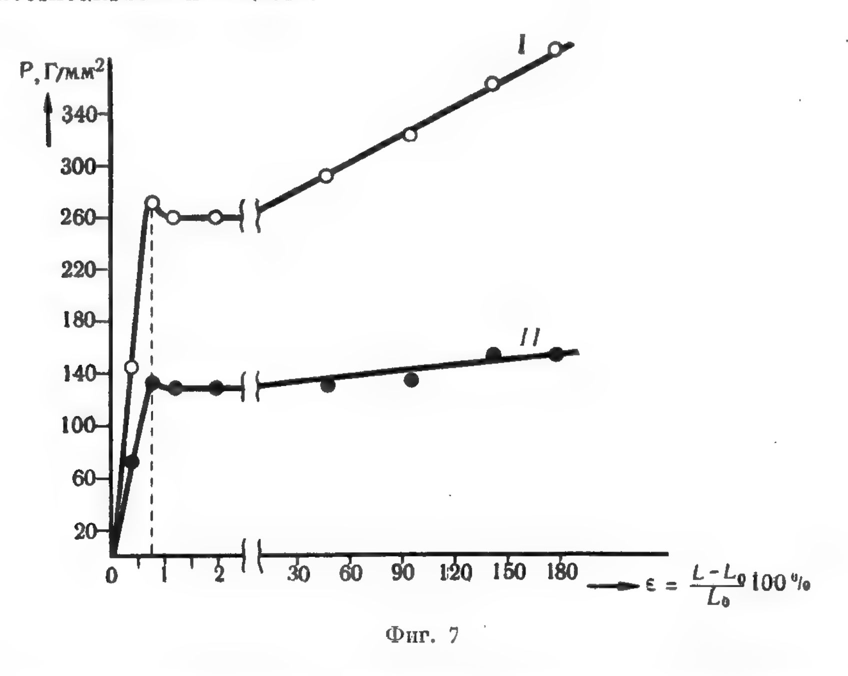
смысле этого слова, причем адсорбционные слои, проникая в микрощели и создавая добавочные раздвигающие усилия, заставляют при действии внешнего деформирующего усилия работать те плоскости скольжения, которые в обычных условиях, т. е. в воздухе или вообще в инактивной среде, не получают развития.



Фиг. 6

На фиг. 7 представлены полные диаграммы растяжения монокристаллов олова в вазелиновом масле с различным содержанием олеиновой кислоты. Эти диаграммы $P = f(\varepsilon)$ прежде всего устанавливают понижение предела текучести почти в два раза под влиянием адсорбирующегося вещества — с 254 до 132 Г/мм² (при 0.2% олеиновой кислоты). С этим связано и соответствующее понижение «модуля» $\left(\frac{dP}{d\varepsilon}\right)_0$, отвечающего начальной стадии пластической деформации. В дальнейшем нами было непосредственно доказано облегчение деформации монокристаллов под влиянием поверхностноактивных веществ и в этой области. Относитель-

ное удлинение ε_0, соответствующее пределу текучести, составляет около 1%, и его величина в этих условиях мало чувствительна к действию поверхностноактивных веществ.



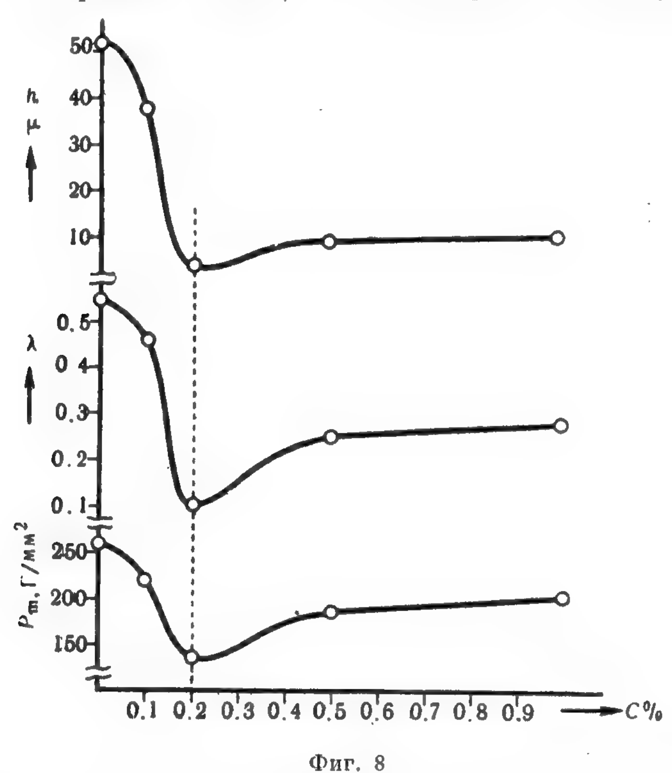
Фиг. 7

В области значительных удлинений действие адсорбирующихся веществ проявляется в значительном понижении коэффициента упрочнения (в 5—6 раз), так как утоньшение пачек скольжения связано — для одинакового удлинения монокристалла — с пропорциональным уменьшением величины сдвига между отдельными пачками.

Все три эффекта действия адсорбирующегося вещества—понижение предела текучести, измельчение пачек скольжения и понижение коэффициента упрочнения — обнаруживают общую зависимость от концентрации олеиновой кислоты (фиг. 8).[1] Относительно наиболее сильные действия адсорбирующихся веществ наблюдаются в области очень малых

[1] Л. П. Янова в нашей лаборатории показала [12], что аналогичные зависимости с тем же порядком величины исследуемых эффектов наблюдаются для пальмитиновой кислоты и для цетилового спирта в качестве поверхностноактивных веществ.

концентраций (около 0.2%), при которых все перечисленные эффекты достигают максимального значения; при дальнейшем повышении концентрации они вновь уменьшаются, что в той или иной степени связано с агрегированием молекул поверхностноактивного вещества в растворе (в углеводородной среде) за счет связи, возникающей между их полярными группами. Интересно отметить, что концентрация максимума эффекта



Фиг. 8

(0.2%) вполне совпадает с соответствующим значением для поликристаллических образцов олова, свинца и меди в наших предыдущих работах [9] и отвечает равновесной концентрации насыщения адсорбционного слоя, зависящей главным образом не от природы металла, а от поверхностноактивного вещества (в соответствии с правилом Траубе эта равновесная концентрация сильно убывает с удлинением углеводородной цепи в гомологическом ряду).

Таблица 1

ВЛИЯНИЕ КОНЦЕНТРАЦИИ ДОБАВКИ ОЛЕИНОВОЙ КИСЛОТЫ К ВАЗЕЛИНОВОМУ МАСЛУ НА СТРУКТУРНО-МЕХАНИЧЕСКИЕ ХАРАКТЕРИСТИКИ МОНОКРИСТАЛЛОВ ОЛОВА, РАСТЯГИВАЕМЫХ С ПОСТОЯННОЙ СКОРОСТЬЮ $V = \frac{d\varepsilon}{d\tau} = 4.8\ \%$ мин.$^{-1}$

Концентрация олеиновой кислоты c % . .	0	0.05	0.10	0.20	0.50	1.00
Предел текучести P_m Г/мм^2 .	254	246	215	132	180	196
Коэф. упрочнения ($\varepsilon = 10$—200%) $\frac{dP}{d\varepsilon}$ . . .	0.56	0.54	0.44	0.10	0.24	0.26
Ср. толщина пачек скольжения h, μ . . .	52	50	38	3.6	8.5	9.4

Пластическое деформирование металлических монокристаллов всегда связано с возникновением и одновременным развитием двух взаимно противоположных процессов: основного, атермического явления — упрочнения, связанного с возникновением внутренних поверхностей раздела, т. е. с изменением характера связей атомов в металле [13] и повышением энергетических уровней атомов металла вдоль этих поверхностей, и вторичного процесса — отдыха, существенно зависящего от температуры и связанного с самопроизвольной перестройкой межатомных связей в сторону понижения свободной энергии системы.

Процесс отдыха в конечном итоге приводит к исчезновению возникающих при упрочнении внутренних поверхностей раздела в деформируемом металле. В области низких температур явление отдыха проявляется в слабой степени и не оказывает заметного влияния на ход пластической деформации металла. С повышением же температуры роль отдыха становится все значительнее, и при достаточно высоких температурах (тем более высоких, чем выше температура плавления данного металла) отдых может полностью затушевать основное явление — процесс упрочнения. В поликристаллических металлах такого рода отдых аналогичен собирательной рекристаллизации, также сводящейся к исчезновению границ раздела между отдельными зернами и представляющей собою процесс роста зерен, протекающий самопроизвольно с уменьшением свободной поверхностной энергии поликристаллической системы. Поверхности раздела, возникаю-

щие при деформации в монокристаллах, в отличие от поликристаллических материалов, носят зародышевый, но вместе с тем и более активный характер. Напомним, что и в поликристаллических металлах рекристаллизация происходит достаточно интенсивно, т. е. с заметной скоростью, только после возникновения в зернах — кристаллитах при достаточной степени пластической деформации таких деформационных повреждений, которые являются зародышами процесса рекристаллизации.

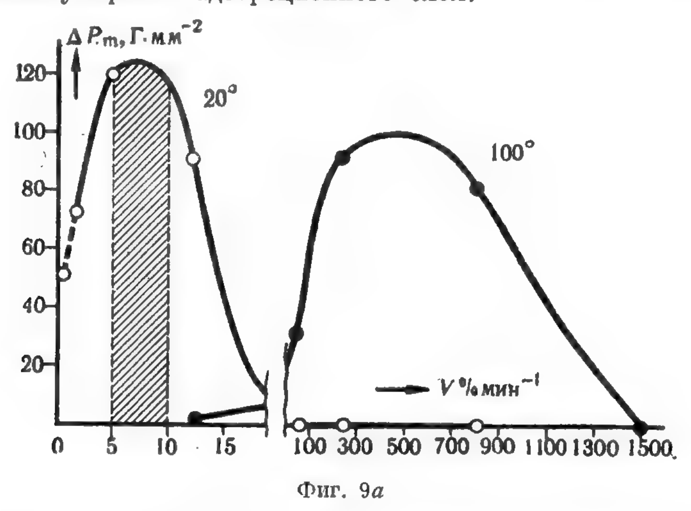
Роль, эквивалентную температуре в установлении определенного соотношения между упрочнением и отдыхом, может играть также скорость деформации. Действительно, понижение скорости деформации, вызывающее усиление процесса отдыха при данной температуре, эквивалентно в этом смысле повышению температуры и наоборот.

Внутренние поверхности раздела, развивающиеся в деформируемом монокристалле вдоль плоскостей скольжения в связи с поворотом и изгибом пачек скольжения, при трансляциях образуют в поверхностном слое кристалла сетку клиновидных микрощелей с устьями, выходящими на внешнюю поверхность, и тупиками, более или менее глубоко проникающими внутрь металла. Возможность адсорбционного воздействия поверхностноактивных веществ, понижающих усилия и работу пластической деформации монокристаллов, определяется развитием именно такой системы микрощелей в поверхностном слое кристалла, как это было показано нами ранее. Исследуя адсорбционные эффекты при различных скоростях деформации монокристаллов, Л. П. Янова [12] дала материал для установления связи между кинетикой развития поверхностной микротрещиноватости в ходе пластической деформации, определяющейся относительным развитием процессов упрочнения и отдыха, и величиной адсорбционного понижения прочности монокристаллов.

В этой работе монокристаллы олова, как уже указывалось, отбирались с оптимальной ориентацией плоскостей скольжения (45°). Для монокристаллов свинца ориентация октаэдрических плоскостей скольжения не столь существенна.

Полные диаграммы растяжения измерялись, как обычно, при помощи аппарата типа Поляни при заданной постоянной скорости деформации (относительного удлинения), причем специально подобранные редукторы позволяли изменять скорость растяжения в широком диапазоне —от 0.05 до 1500% мин.$^{-1}$, т. е. в отношении 1:30 000. Растяжение проводилось в

воздухе или неполярном вазелиновом масле глубокой очистки (что давало всегда одинаковые результаты) и в вазелиновом масле с добавками к нему в качестве поверхностноактивных веществ олеиновой и пальмитиновой кислот и цетилового спирта в оптимальных концентрациях (около 0.2%). В предыдущих работах было установлено, что эта концентрация соответствует максимуму эффекта и образованию насыщенного мономолекулярного адсорбционного слоя.

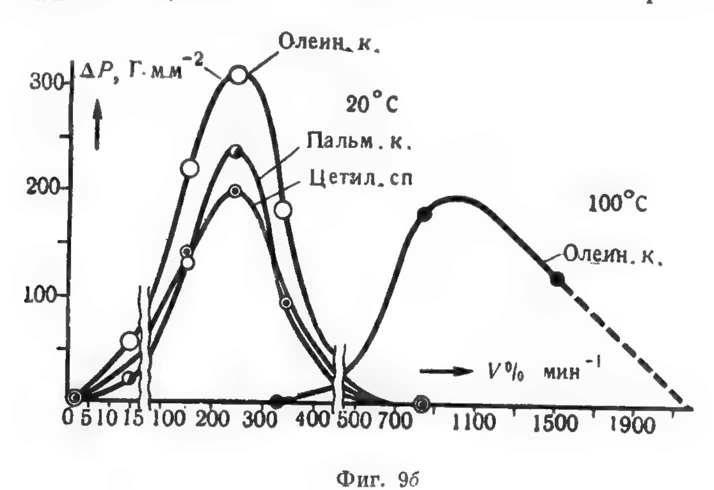


Фиг. 9*а*

На фиг. 9*а* представлены результаты измерений величины адсорбционного понижения прочности монокристаллов олова, а на фиг. 9*б* — свинца от скорости деформации при двух температурах: 20 и 100°С. За меру адсорбционного эффекта (отложенного по оси ординат) $\Delta P = P_0 - P_a$ для монокристаллов олова принимается, как обычно, относительное понижение предела текучести под влиянием добавки поверхностноактивного вещества в оптимальной концентрации (P_0 — предел текучести в инактивной среде — воздухе или неполярном вазелиновом масле, а P_a — в наиболее поверхностноактивной среде). Для монокристаллов свинца, не имеющих явно выраженного предела текучести, он заменяется в расчетах наибольшим значением напряжения, отмечаемым на диаграмме. Для олова при 20° максимум адсорбционного эффекта лежит при скорости деформации $V \sim 5\%$ мин.$^{-1}$ При уменьшении скорости деформации до 0.05% мин.$^{-1}$ адсорбционный эффект практически исчезает и не наблю-

дается более при дальнейшем замедлении деформации. По другую сторону максимума, при повышении скорости деформации, адсорбционный эффект практически исчезает при V около 150% мин.$^{-1}$

При повышенной температуре (100°С) для олова максимум адсорбционного эффекта переходит к значительно большим скоростям — около



Фиг. 9б

300% мин.$^{-1}$, практически исчезая при понижении скорости до 15%мин.$^{-1}$ и при повышении скорости до величин порядка 1000% мин.$^{-1}$

Для свинца установлена в общем та же зависимость со специфически иными, конечно, значениями скоростей деформации V_m, отвечающими максимуму адсорбционного эффекта (фиг. 9б).

Установленные результаты определяют исключительно важную роль процесса отдыха металлического кристалла в явлениях адсорбционного воздействия на его пластическую деформацию; если скорость деформации такова, что процесс отдыха успевает почти полностью произойти за время деформации, то соответствующее этому смыкание поверхностных микрощелей в самом процессе деформации будет препятствовать воздействию адсорбирующихся веществ. Только при не очень малых скоростях деформации, когда явлениями отдыха за время деформации можно пренебречь, адсорбционные воздействия осуществляются в полной мере. Именно таким скоростям и соответствует область максимума адсорбционного эффекта.

Если же скорости деформации становятся столь значительными, что процесс принимает ударный характер, то при этом, как известно, нарушается простой трансляционный механизм пластической деформации монокристалла, и в строй вступают сразу несколько плоскостей скольжения. Это связано с понижением пластических свойств металла, с его разрушением хрупкого характера при сравнительно малых деформациях. Микрощели, возникающие в этом случае на поверхности образца, не успевают достаточно развиться в глубь металла, и деформация, становясь неравномерной, сосредоточивается в одном наиболее слабом месте кристалла, где она и заканчивается разрывом. При этом скорость образования новой поверхности в микрощелях слабого участка столь велика, что эта поверхность не успевает покрываться адсорбционным слоем.

Повышение температуры, ускоряя процесс отдыха металлического кристалла, смещает максимум адсорбционного эффекта в сторону больших скоростей деформации в соответствии с уже указанной схемой. При этом некоторое уменьшение величины адсорбционного эффекта понижения прочности монокристаллов при повышении температуры может быть связано с изменением характера адсорбции в сторону перехода к активированной (химической) адсорбции и с уменьшением подвижности адсорбированных молекул.

Полученные результаты показывают, что величина адсорбционного эффекта облегчения деформаций (понижения прочности) кристаллов определяется режимом деформации — скоростью ее протекания и температурой — и может наблюдаться только в определенной, хотя и достаточно широкой области некоторых средних оптимальных скоростей деформации. Положение этой области определяется температурой измерения.

Изучение упругих свойств металлических монокристаллов [1] показало, что величина упругого сдвига в них исчезающе мала. Таким образом, недеформированный монокристалл в первом приближении может рассматриваться как идеально пластическое тело в том смысле, что он вовсе лишен упругой области и что предел текучести соответствует деформациям (удлинениям), весьма малым по сравнению с их возможными значениями в области пластического течения. Эта область наступает тогда, когда напряжение достигает некоторой критической величины в действующей системе скольжения. Но по мере развития пластической деформации

и связанного с ней расчленения однородного кристалла на отдельные блоки с образованием внутренних поверхностей раздела в результате поворотов и изгиба пачек скольжения, можно ожидать значительного развития упругой области по аналогии с тем, как повышение дисперсности (измельчение зерна) расширяет область упругих деформаций в поликристаллах.

Однако сразу же необходимо указать на принципиальное различие этих явлений для монокристаллов и для поликристаллических агрегатов. В поликристаллических металлах беспорядочная ориентировка зерен приводит к различной устойчивости их по отношению к внешним деформирующим усилиям. Зерна, расположенные благоприятно к направлению этих усилий, могут подвергнуться пластической деформации, тогда как неблагоприятно расположенные зерна еще останутся в области чисто упругой деформации. Поэтому на границах зерен создаются объемно-напряженные области, затрудняющие дальнейшее развитие пластической деформации и способствующие расширению упругой зоны. При увеличении дисперсности поликристаллического агрегата, т. е. при переходе к мелкозернистым структурам, возрастает значение объемно-напряженных участков на границах зерен в общей системе внутренних напряжений в каждом зерне и во всем образце в целом, что вызывает развитие упругих свойств.

В монокристаллах же ничто не тормозит свободного развития трансляционного механизма соскальзывания отдельных блоков решетки, и они сохраняют в этом процессе единую кристаллографическую ориентировку. Улучшение упругих свойств при деформации связано при этом с увеличением избытка свободной энергии монокристалла в процессе деформации, главным образом за счет возрастания поверхностной энергии вдоль плоскостей скольжения при образовании и развитии клиновидных микрощелей, а также за счет упругого изгиба пачек скольжения. Первый фактор особенно важен для влияния адсорбции из внешней среды на процессы упругой и пластической деформации монокристаллов, так как именно внутренние поверхности раздела, зарождающиеся и развивающиеся при деформации, являются проводниками этого воздействия.

Значительное развитие упругой области в деформированных монокристаллах было обнаружено нами при изучении процесса «отдыха» предварительно растянутых монокристаллов олова высокой чистоты после сня-

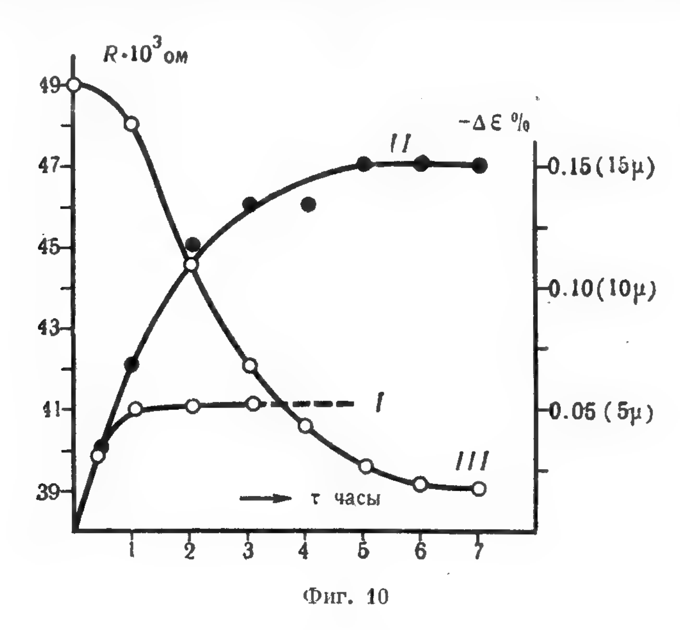
тия нагрузки. Монокристаллы в виде проволоки диаметром около 1 мм подвергались растяжению на приборе типа Поляни с постоянной скоростью удлинения. Опыты велись в инактивной среде — чистом вазелиновом масле и в вазелиновом масле с добавкой к нему 0.2% олеиновой кислоты в качестве поверхностноактивного вещества. В процессе растяжения, кроме снятия полной диаграммы деформации $P = f(\varepsilon)$, измерялась также электропроводность образца. После того как достигалась заданная степень растяжения (около 200%), образец разгружался путем обратного вращения винта до нулевого положения оптического динамометра. С этого момента и в течение довольно продолжительного времени, пока длилась первая стадия процесса отдыха (около 45 час.), нами наблюдалось постепенное сокращение длины образца, выражавшееся в возникновении усилия, действующего на оптический динамометр прибора; для снятия этого усилия требовалось дальнейшее вращение винта.

Общая величина такого своеобразного явления упругого последействия разгрузки в инактивной среде — воздухе или чистом вазелиновом масле — достигала 10—15 μ, или около 0.03% длины образца (при точности отсчета 1—5 μ) и требовала около 1.5 часа. В присутствии поверхностноактивных добавок к углеводородной среде величина упругого последействия возрастала до 40—50 μ, или до 0.1—0.15%, и время завершения этого явления увеличивалось до 6—7 час. Одновременно с развитием упругого последействия происходило некоторое возрастание электропроводности деформированных монокристаллов, достигавшее 3—4% в инактивной среде и 15—20% в активной среде, где конечное значение электропроводности при той же степени удлинения значительно ниже, чем в инактивной среде [3]. На фиг. 10 представлены ход упругого последействия в разных средах и кривая падения электрического сопротивления со временем для монокристаллов, деформированных в активной среде.

В работах Е. Шмида и его сотрудников, подробно изучавших процесс отдыха предварительно деформированных металлических монокристаллов после снятия нагрузки, установлено, что полное восстановление начальных механических свойств (оцениваемых по величине предела текучести) при комнатной температуре для олова происходит, примерно, в течение суток и зависит от температуры и степени упрочнения.

Таким образом, обнаруженное явление упругого последействия, являясь составной частью отдыха деформированного кристалла, не исчерпы-

вает собою этого процесса полностью. Весь процесс отдыха может быть разделен на три главных стадии. Первая, самая кратковременная, полностью завершается уже при снятии нагрузки (со скоростью звука) и связана с частичным исчезновением «истинных» упругих напряжений в кристалле. Вторая, более продолжительная, стадия состоит в постепенном исчезно-



Фиг. 10

вении внутренних поверхностей раздела, возникающих в процессе деформации вдоль плоскостей скольжения, и в смыкании возникших на их основе в поверхностном слое клиновидных микрощелей. И, наконец, последняя, наиболее продолжительная, третья стадия включает в себя дальнейшие постепенные уменьшения свободной поверхностной энергии в местах «повреждений решетки», связанные с восстановлением нарушенных при деформации межатомных связей с их частичной перестройкой,— своеобразной рекристаллизацией.

Упругое последействие разгрузки связано главным образом со второй стадией процесса отдыха, являясь результатом сил молекулярного сцепления, действующих в решетке у внутренних концов (в тупиках микрощелей) и вызывающих их смыкание. Значение адсорбции в этом явлении

сводится, повидимому, к развитию роли второй, медленной, стадии за счет первой, быстрой в общем процессе отдыха, а может быть, и за счет сокращения третьей стадии вследствие меньшего повреждения решетки при деформации в присутствии поверхностноактивных веществ. Действительно, как это было нами показано, адсорбционные действия поверхностноактивных веществ приводят к значительному снижению упрочнения деформируемого монокристалла [11] и, следовательно, к сравнительно малой порче решетки. Кроме того, адсорбционные слои, вызывая значительное увеличение числа работающих пачек скольжения и проникая внутрь металла, уменьшают истинные упругие напряжения в кристалле, переводя часть упругой энергии деформируемого кристалла в свободную поверхностную энергию развивающихся микрощелей. Это делает понятным возрастание величины упругого последействия в присутствии адсорбирующихся веществ. Последействие связано при этом со смыканием гораздо более глубоких микрощелей. Адсорбционные слои, покрывающие стенки микрощелей, должны быть вытеснены наружу, что замедляет процесс смыкания, который можно проследить и по нарастанию электропроводности монокристалла при отдыхе.

Мы уже указывали на значительную роль поверхностной микротрещиноватости при отдыхе, протекающем в самом процессе пластической деформации. Явление упругого последействия разгрузки деформированных монокристаллов показывает, что роль поверхностных микрощелей не менее велика и в процессах отдыха после разгрузки. В последнее время закономерности образования и развития микрощелей на поверхности монокристалла в процессе деформации, особенно в присутствии адсорбирующихся веществ, прослежены нами с помощью электронно-микроскопического изучения поверхности алюминиевых монокристаллов и поликристаллического алюминия.

Результаты, изложенные в этой статье, показывают, сколь велико значение ранее не изучавшихся физико-химических особенностей процессов деформации для физики твердого тела, металловедения и учения о механической обработке материалов.

Перед этой областью физико-химической механики, возникшей в исследованиях советских ученых, — широкие перспективы дальнейшего развития и плодотворного использования в различных прикладных проблемах.

ЛИТЕРАТУРА

1. *П. А. Ребиндер.* ЖТФ, т. 2, № 7—8, 1932, стр. 726; ЖФХ, т. 5, № 2—3, 1934 Zs. f. Phys., т. 72, 1931, стр. 91.
2. *П. А. Ребиндер.* Новые физико-химические явления в процессах деформации и обработки твердых тел (доклад на Общем собрании АН СССР 2/VII 1946), Труды июльской сессии АН СССР, Изд. АН СССР, М. — Л., 1946.
3. *Б. В. Дерягин* и *М. М. Кусаков.* Изв. АН СССР, ОМЕН, серия хим., № 5, 1936, стр. 741.
4. *П. А. Ребиндер* и *Г. И. Логгинов.* ДАН СССР, нов. серия, т. 31, № 6, 1941, стр. 489.
5. *П. А. Ребиндер, Л. А. Шрейнер* и *К. Ф. Жигач.* Понизители твердости в бурении (монография). Изд. АН СССР, М. — Л., 1944; *П. А. Ребиндер.* Вестн. АН СССР, № 8—9, 1940, стр. 5.
6. *И. В. Обреимов.* Proc. Roy. Soc. (A), т. 127, 1930, стр. 290.
7. *А. Ф. Иоффе.* Изв. СПб. Политехн. ин-та, 1906.
8. *E. Orowan.* Nature, № 152, 1944.
9. *П. А. Ребиндер* и *Е. К. Венстрем.* Изв. АН СССР, ОМЕН, серия физ., № 4—5, 1937, стр. 531; *П. А. Ребиндер.* Изв. АН СССР, ОМЕН, серия хим., т. 5, 1936, стр. 636.
10. *П. А. Ребиндер, В. И. Лихтман, Б. М. Масленников.* ДАН СССР, нов. серия, т. 32, № 2, 1941, стр. 127; *В. И. Лихтман* и *П. А. Ребиндер.* ДАН СССР, нов. серия, т. 32, № 2, 1941, стр. 131.
11. *П. А. Ребиндер* и *В. И. Лихтман.* ДАН СССР, нов. серия, т. 56, 1947, № 7, стр. 723.
12. *В. И. Лихтман, П. А. Ребиндер* и *Л. П. Янова.* ДАН СССР, нов. серия, т. 56, 1947, № 8, стр. 827.
13. *Г. В. Акимов.* ДАН СССР, нов. серия, т. 47, № 8, 1945, стр. 567.

Академик
М. М. Дубинин

СТРУКТУРА И СОРБЦИОННЫЕ СВОЙСТВА АКТИВНЫХ УГЛЕЙ

В современной технике обширное и разнообразное применение имеют пористые адсорбенты, и среди них первое место, несомненно, принадлежит активным углям. Это особое положение активных углей в значительной мере определяется гидрофобностью угольной поверхности, или относительно малым ее сродством к воде. На практике наибольшее значение имеет поглощение органических паров из воздуха, всегда содержащего водяные пары, или из водных растворов. Поэтому из смеси органических паров и паров воды или из водных растворов активные угли адсорбируют преимущественно посторонние примеси или растворенные вещества.

В больших количествах активные угли применяются в процессах поглощения, осуществляемых в газовой фазе, например для улавливания и обратного возвращения на производство паров ценных органических растворителей, выделения бензина из природных газов, тонкой очистки газов для каталитического синтеза, очищения воздуха от отравляющих веществ в противогазовом деле и т. п.

Не менее важное значение имеют и процессы поглощения в жидкой фазе для удаления примесей из жидкостей и водных растворов или для выделения растворенных веществ. Активные угли применяются для обесцвечивания сахарных сиропов, удаления сивушных масел из спирта, очищения питьевой воды, а также промышленных и сточных вод и во

многих других случаях. Активный уголь зачастую является незаменимым адсорбентом для выделения растворенных веществ из разбавленных водных растворов, например в производстве пенициллина.

В ряде процессов активные угли применяются в качестве катализаторов или носителей для каталитически активных веществ. Например, сероводород как примесь к светильному или природному газу легко окисляется на поверхности активного угля кислородом добавленного воздуха до элементарной серы. Так осуществляется не только очистка газов от сернистых соединений, но и получаются значительные количества достаточно чистой серы.

В зависимости от исходных материалов, метода получения и последующей обработки активные угли обычно содержат до 10% зольных неорганических веществ. Органическое вещество активного угля в основном состоит из углерода (95—98%). Почти весь остаток приходится на долю водорода и кислорода, преимущественно связанных с поверхностными атомами углерода в угле.

Рентгенографическими исследованиями доказана кристаллическая структура углерода в активных углях [1]. Элементарные кристаллиты углерода в активных углях имеют решетку графита. Размеры кристаллитов зависят от температуры выделения углерода в процессе термического разложения исходных органических соединений; в среднем их диаметр составляет около 23Å и высота около 9Å. Цилиндрическим кристалликам таких размеров отвечает суммарная поверхность около 1800 $м^2$ для 1 г углерода. Доступная для адсорбции поверхность 1 г активных углей оценивается в 400—700 $м^2$/г (удельная поверхность угля). Таким образом, активный уголь можно рассматривать как пористое тело, состоящее из совокупности элементарных кристаллов графита. В зависимости от общего объема и размеров пор кристаллиты находятся в более или менее тесном соприкосновении друг с другом, что и определяет величину удельной поверхности, имеющей значение для адсорбции.

ПОРИСТОСТЬ АКТИВНЫХ УГЛЕЙ

Пористость активных углей — основной элемент их структуры, определяющий сорбционные свойства. Размеры пор активных углей заключаются в широком интервале радиусов и простираются от пор, видимых

в лупу, до пор молекулярных размеров. Именно наличие очень мелких пор приводит к большой удельной поверхности активных углей порядка сотен квадратных метров на грамм. В общем случае характер пористости угля зависит от распределения объема и поверхности пор по их радиусам.

Степень развития пористости активного угля характеризуется суммарным объемом пор v_{ε}, выражающимся разностью объема, занимаемого 1 г зерен угля, и истинным объемом собственного вещества угля. Эти объемы легко вычисляются по кажущемуся δ и истинному d удельным весам угля

$$v_{\varepsilon}=\frac{1}{\delta}-\frac{1}{d}. \quad (1)$$

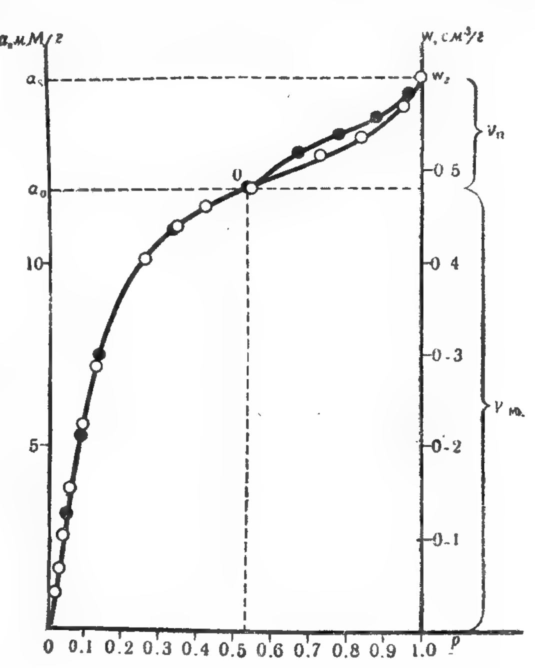
Это, однако, только самая общая оценка пористости, так сказать, первый этап в описании свойств активного угля. Второй этап, детализирующий структуру угля, состоит в определении суммарных объемов характерных типов пор: микропор, переходных пор и макропор. Микропоры — самые мелкие поры угля, с размерами порядка молекулярных. Как будет показано ниже, этим порам практически принадлежит вся поверхность, на которой происходит адсорбция. Переходные поры относятся к более крупной разновидности; их максимальные размеры не превышают десятитысячной доли сантиметра. При сорбции паров, например метилового спирта, в области высоких относительных давлений, вплоть до давления насыщения, в этих порах происходит капиллярная конденсация, т. е. ожижение паров в результате пониженной упругости насыщенного пара над вогнутыми менисками жидкости. Наконец, более крупные поры, включая и непосредственно видимые при небольшом увеличении, относятся к макропорам угля.

Методы определения объема микропор были разработаны в последние годы [2—4]. Один из методов заключается в измерении изотерм сорбции и десорбции паров метилового спирта при 20° С [3]. Типичная форма изотермы для активного угля изображена на фиг. 1. Ее обратимая область отвечает адсорбции; при относительно высоких равновесных давлениях происходит капиллярная конденсация, для которой характерно несовпадение ветвей сорбции и десорбции изотермы (явление гистерезиса). Точке O отвечает заполнение адсорбирующих пор, или микропор, активного угля ожиженным паром. Произведение величины адсорбции

a_0 для точки начала гистерезиса O на молярный объем жидкости v непосредственно выражает объем микропор угля (на правой вертикальной оси фиг. 1 отложены объемы ожиженного пара w в порах угля)

$$v_{\text{ми}} = a_0 v. \tag{2}$$

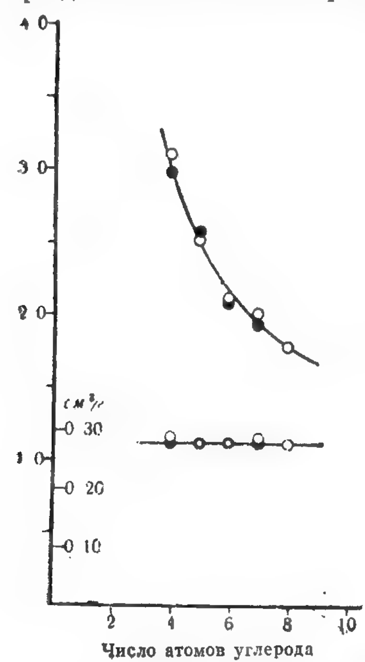
Для многих изученных образцов углей точке O изотермы отвечает относительное давление паров метилового спирта 0.53 при 20° С. Так как капиллярная конденсация происходит в более крупных переходных порах угля, то, по теории капиллярной конденсации, легко вычислить относительное давление другого пара, например бензола, при котором объемы ожиженных паров w метилового спирта и бензола равны [1]. Это, соответственное точке O, относительное давление при 20° С составляет для бензола 0.17. Произведение величины адсорбции бензола при этом равновесном давлении на молярный объем его, согласно (2), также выражает объем микропор. Это значительно упрощает определение, так как бензол является стандартным паром, для которого обычно измеряются изотермы адсорбции.



Фиг. 1. Изотерма сорбции паров метилового спирта при 20°С активным углем

Другой метод заключается в определении предельно адсорбированных активными углями объемов жирных кислот или спиртов из водных растворов [2]. Для непористых адсорбентов, а также некоторых обесцвечивающих углей, в которых очень мелкие поры практически отсут-

ствуют [4], предельная адсорбция, выраженная в молях и достигаемая при достаточно больших равновесных концентрациях, не зависит от длины углеводородной цепи молекул. Это свидетельствует о мономолекулярной адсорбции при насыщении на постоянной по величине поверхности угля. Для относительно мелкопористых углей, применяемых для адсорбции парообразных веществ, предельные величины адсорбции уменьшаются с ростом углеводородного радикала. Однако произведения их на молярные объемы веществ в жидком состоянии оказываются постоянными для каждого образца угля (правило постоянства предельно адсорбированных объемов) и практически равными объему микропор, определенному по первому методу. Эти соотношения [5] наглядно иллюстрируются опытными данными (фиг. 2).



Фиг. 2. Предельно адсорбированные количества (верхняя кривая — a, мМ/г) и соответствующие объемы (нижняя кривая — av, см³/г) нормальных жирных кислот (белые кружки) и спиртов (черные кружки) в зависимости от числа атомов углерода в молекулах для адсорбции из водных растворов на активном угле

Таким образом, вблизи области насыщения молекулы растворенных веществ «выталкиваются» из водных растворов в тонкие поры угля с более высокими адсорбционными потенциалами и заполняют их с плотностями, приближающимися к значениям для чистых веществ. В табл. 1 приведены объемы микропор для различных активных углей, определенные рассмотренными методами.

В достаточно удовлетворительном приближении эти методы приводят к одинаковым значениям объемов микропор для каждого угля. Методы, основанные на измерениях адсорбции парообразных и растворенных веществ, независимы. Поэтому полученные результаты служат подтверждением реальности физического образа адсорбирующих пор, или микропор, активных углей.

Таблица 1

***ОБЪЕМЫ МИКРОПОР АКТИВНЫХ УГЛЕЙ**, см³/г*

Вычислено по измерениям адсорбции	*Образцы углей*				
	1	*2*	*3*	*4*	*5*
Жирных кислот и спиртов из водных растворов	0.27	0.26	0.38	0.31	0.59
Паров метилового спирта	0.26	0.24	0.39	0.31	0.60
Паров бензола	0.25	0.23	0.38	0.33	0.60

Переходные или конденсирующие поры активных углей заполняются в процессе капиллярной конденсации паров метилового спирта. Их объем $v_п$ выражается разностью максимального сорбированного объема пара в виде жидкости $w_s = a_s v$ и объема микропор угля

$$v_п = w_s - v_{ми}. \qquad (3)$$

Более крупные поры угля относятся к классу макропор. Даже самые мелкие макропоры, с радиусами порядка десятитысячной доли сантиметра, не заполняются при капиллярной конденсации, так как понижение упругости насыщенного пара метилового спирта над менисками жидкости в этих порах составляет около 0.1% и ничтожные примеси даже в химически чистом препарате спирта приводят к снижению давления насыщенного пара метилового спирта того же порядка. Макропоры могут быть заполнены только при пропитке угля жидкостью. Их объем выражается разностью суммарного объема пор угля $v_в$ [формула (1)] и максимального сорбированного объема w_s

$$v_{ма} = v_в - w_s. \qquad (4)$$

Порядок суммарной поверхности, соответствующей каждой разновидности пор угля, может быть оценен из простых геометрических соображений. Представляя поры с объемом v в виде цилиндров со средним радиусом $\bar{r}$, получим для внутренней поверхности пор s выражение

$$s = \frac{2\bar{v}}{\bar{r}} \left[\frac{см^2}{г}\right]. \qquad (5)$$

Объем микропор различных активных углей, применяющихся для адсорбции парообразных веществ, заключается в пределах 0.25—0.60 $см^3$/г. Их средний радиус может быть оценен приблизительно в $1.4 \cdot 10^{-7}$ см. Согласно (5), общая поверхность микропор для 1 г активных углей изменяется в пределах 350—850 $м^2$/г.

Объем переходных пор изменяется в пределах 0.05—0.15 $см^3$/г. Их средний радиус составляет приблизительно $5 \cdot 10^{-6}$ см. Суммарная поверхность переходных пор активных углей заключается в пределах 1—3 $м^2$/г.

Объем макропор углей лежит в пределах 0.15—2.5 $см^3$/г и обычно составляет 0.35—0.60 $см^3$/г. Средний радиус этого типа пор приблизительно равен $1 \cdot 10^{-3}$ см. Суммарная поверхность макропор угля имеет порядок десятых долей 1 $м^2$/г.

Таким образом, вся удельная поверхность активного угля, имеющая значение для адсорбции, практически целиком принадлежит микропорам угля. Поэтому при вычислении объема микропор из опытов адсорбции паров возможно пренебречь количеством адсорбированного пара на стенках более крупных пор угля.

Объем микропор служит только общей характеристикой развития адсорбирующих пор в активном угле. Адсорбционные свойства угля определяются характером пористости для этого типа пор, т. е. распределением объема и поверхности пор по радиусам. Мы не располагаем непосредственными методами исследования структуры микропор активных углей. Поэтому о содержании в угле адсорбирующих пор различных размеров приходится судить по результатам косвенных опытов.

Размеры микропор угля имеют важное значение для физической адсорбции парообразных и растворенных веществ [1]. С одной стороны, в порах или щелях угля молекулярных размеров происходит значительное увеличение адсорбционных потенциалов или энергии адсорбции в результате наложения полей противоположных стенок [6]. Этот эффект приводит к резкому возрастанию адсорбции, в особенности в области относительно малых равновесных концентраций. С другой стороны, поверхность активного угля, имеющая значение для адсорбции, зависит от размеров адсорбируемых молекул, так как часть более мелких микропор угля оказывается недоступной для молекул. Этот эффект сказывается тем резче, чем мельче поры угля и чем крупнее адсорбируемые молекулы.

Для активных углей, как и вообще для пористых адсорбентов, вели-

чины адсорбции принято относить к единице массы угля (a, мМ/г). Если α — величина адсорбции для единицы поверхности адсорбента (α, мМ/см^2) и s — поверхность угля, доступная для рассматриваемых молекул (s, см^2/г), то

$$a = \alpha s. \tag{6}$$

Для крупнопористых активных углей, где удельная поверхность s практически не зависит от размеров адсорбируемых молекул, обычно наблюдается возрастание адсорбции по мере перехода от низших к высшим членам гомологических рядов органических соединений. Так, при адсорбции из водных растворов представителей гомологических рядов одноосновных жирных кислот и спиртов величины поглощения закономерно увеличиваются с возрастанием числа атомов углерода в молекулах (ряд Траубе). Аналогичная картина наблюдается для ряда неорганических кислот, соляной, серной и фосфорной, по мере увеличения молекулярного веса. Такая последовательность в адсорбируемости характерна для процесса адсорбции, происходящего собственно на поверхности раздела уголь — водный раствор, т. е. для значений α, так как в рассматриваемом случае в выражении (6) s = const и a пропорционально α.

Для мелкопористых углей с преобладанием микропор молекулярных размеров, например для углей из сахара, влияние уменьшения адсорбирующей поверхности s с ростом размеров молекул может превысить возрастание величин адсорбции в гомологических рядах или в ряду неорганических кислот. Тогда наблюдается обращение адсорбционных рядов, свидетельствующее о чрезвычайно мелких микропорах угля [7, 8]. По мере развития пористости и общего возрастания размеров пор адсорбционные ряды становятся сперва смешанными, а потом прямыми (ряд Траубе). Последнее свидетельствует о том, что для адсорбируемых веществ с не слишком крупными молекулами удельная поверхность угля почти перестает зависеть от их размеров. Это позволяет сделать определенные выводы о размерах преобладающих в угле пор. В табл. 2 приведены данные об адсорбционных свойствах углей из сахара, расположенных в порядке возрастания размеров микропор [7]. Измерения адсорбции производились из разбавленных водных растворов; величины адсорбции выражены в относительных единицах.

Для мелкопористого угля 1 наблюдается обращение адсорбционных рядов органических и неорганических кислот. Наиболее крупнопористый

Таблица 2

РЯДЫ АДСОРБЦИИ ОРГАНИЧЕСКИХ И НЕОРГАНИЧЕСКИХ КИСЛОТ

Кислота	*Уголь 1*	*Уголь 2*	*Уголь 3*
Пропионовая	21	52	45
Валерьяновая	14	76	75
Гептиловая	5	65	92
Соляная	26	51	48
Серная	8	56	55
Фосфорная	5	84	91

уголь из сравниваемых образцов характеризуется прямыми адсорбционными рядами. Смешанный ряд органических кислот для угля 2 свидетельствует о промежуточном положении этого образца по размерам пор. Для мелкопористых углей типа 1 состояние равновесия при адсорбции из растворов устанавливается значительно медленнее [9]. Это позволяет также предполагать наличие довольно большого количества мелких пор в активных углях.

Увеличивая размеры адсорбируемых из растворов молекул или частиц, можно применить рассмотренный относительный метод характеристики пористости для образцов углей с более крупными микропорами. В табл. 3 приведены опытные данные по адсорбции из водных растворов органических кислот, фенола ($M = 94$), метиленовой голубой ($M = 374$) и конго-рот (субмикроскопические коллоидные частицы) для крупнопористых активных углей из сахара [8]. Как и раньше, величины адсорбции выражены в относительных единицах.

Для всех образцов углей наблюдается нормальный ряд Траубе при адсорбции органических кислот жирного ряда. Несущественные различия в адсорбции кислот свидетельствуют о близких величинах поверхностей активных углей, на которых происходит адсорбция рассматриваемых молекул. Этот вывод относится также и к адсорбции фенола. Однако поверхности пор углей, доступных для адсорбции более крупных частиц красителей, резко отличаются. Так, уголь А5 не содержит в за-

метных количествах пор, в которые могут проникнуть субмикроскопические коллоидные частицы конго-рот; значительная часть пор этого угля недоступна и для молекул метиленовой голубой. Для угля А7 бо̀льшая часть пор доступна для адсорбции наиболее крупных частиц. Уголь А6 занимает промежуточное положение.

Таблица 3

ВЕЛИЧИНЫ АДСОРБЦИИ РАСТВОРЕННЫХ ВЕЩЕСТВ

Адсорбируемое вещество	*Уголь А5*	*Уголь А6*	*Уголь А7*
Пропионовая кислота	43.2	40.8	50.4
Валерьяновая »	80.5	88.3	89.6
Гептиловая »	96.4	98.0	98.8
Фенол	85.4	91.9	93.5
Метиленовая голубая	57.5	96.7	100.0
Конго-рот	0.0	30.0	99.9

Рассмотренные измерения адсорбции растворенных веществ с различными размерами частиц позволяют дать относительную характеристику пористости активных углей и расположить активные угли в ряд по возрастающим размерам микропор. Объемы микропор непосредственно связаны с развитием удельной поверхности активных углей; объемы переходных пор и макропор являются важной характеристикой углей как носителей для катализаторов и хемосорбентов.

ХАРАКТЕР ПОРИСТОСТИ И СОРБЦИОННЫЕ СВОЙСТВА АКТИВНЫХ УГЛЕЙ

Наиболее важное значение имеет применение активных углей для сорбции парообразных веществ. Поэтому мы ограничимся обсуждением этого основного случая сорбции.

Обычно при сорбции парообразных веществ отсутствует химическое взаимодействие, и тепловые эффекты сорбции не превышают 10—15 тыс. кал/моль. Это свидетельствует о том, что сорбция есть результат проявления молекулярных сил, обусловливающих и другие аналогичные яв-

ления, например конденсацию паров в жидкость. С другой стороны, сорбция в области высоких давлений, приближающихся к упругости насыщенного пара p_s, не представляет практического интереса. Поэтому в очень широкой области относительных давлений, до $\frac{p}{p_s}=0.5\div0.8$, сорбция паров сводится к физической адсорбции, т. е. к уплотнению молекул пара на поверхности активного угля, образованию сжатого слоя пара и конденсации пара в жидкую пленку, заполняющую мелкие поры угля [6].

Во взаимодействии молекул пара и поверхности активного угля, представляющего собою аполярное вещество, основную роль играет дисперсионная составляющая. Для элемента плоской поверхности сила притяжения молекул обратно пропорциональна четвертой степени расстояния от поверхности, и адсорбционный потенциал, или работа, совершаемая адсорбционными силами на перемещение молекулы из бесконечности на данное расстояние от поверхности адсорбента, обратно пропорционален третьей степени расстояния. Обычно адсорбционный потенциал относится к одному молю пара.

В мелких порах или щелях угля с размерами порядка молекулярных происходит наложение полей противоположных стенок пор, что приводит к возрастанию адсорбционных потенциалов в несколько раз [6]. Следовательно, в энергетическом отношении поверхность активных углей не является равноценной; наиболее активные участки соответствуют мелким порам. В области малых относительных давлений описанная схема уплотнения молекул пара и образование конденсированных пленок будут преимущественно осуществляться в подобных мелких порах. Таким образом, величины адсорбционных потенциалов тесно связаны со структурой активных углей.

Каждой величине адсорбции пара a при равновесном давлении p отвечает определенный объем конденсированной пленки w на поверхности и в порах угля

$$w=av, \tag{7}$$

где v — молярный объем жидкости. Из теории дисперсионных сил следует, что силы взаимодействия между молекулами в конденсированной пленке не зависят от поля, создаваемого поверхностью адсорбента [6]. Поэтому поверхности жидкой пленки соответствует нормальное значение

упругости насыщенного пара p_s при температуре опыта, и поверхность пленки будет эквипотенциальной, т. е. отвечать одинаковым значениям адсорбционных потенциалов.

Работа переноса моля пара из объема сосуда на поверхность конденсированной пленки, т. е. величина адсорбционного потенциала для ее поверхности ε, равна изотермической работе сжатия пара от равновесного давления p до давления насыщения p [6]

$$\varepsilon = RT \ln \frac{p_s}{p}. \qquad (8)$$

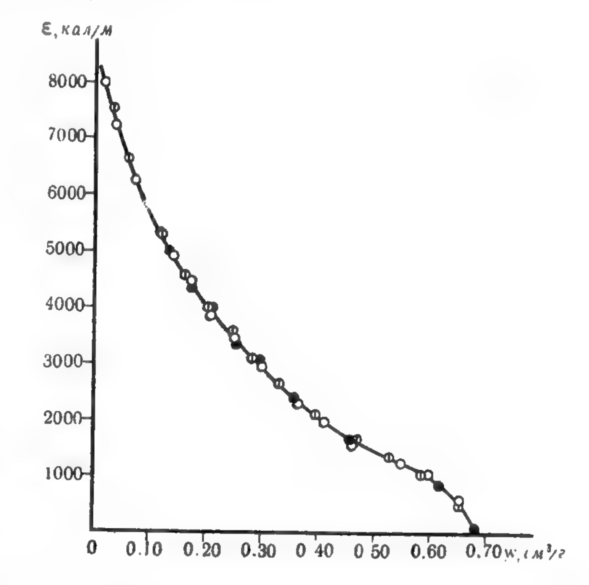
Поэтому изотерме адсорбции пара

$$a = f(p) \qquad (9)$$

соответствует, в связи с (7) и (8), уравнение характеристической кривой

$$\varepsilon = \varphi(w), \qquad (10)$$

выражающее зависимость адсорбционного потенциала ε от объема конденсированной пленки w. Уравнение характеристической кривой не зависит от температуры, так как на дисперсионные силы и, следовательно, на их работу ε температура не оказывает влияния. Объемом же конденсированной пленки w задается только определенная эквипотенциальная поверхность поля адсорбента.



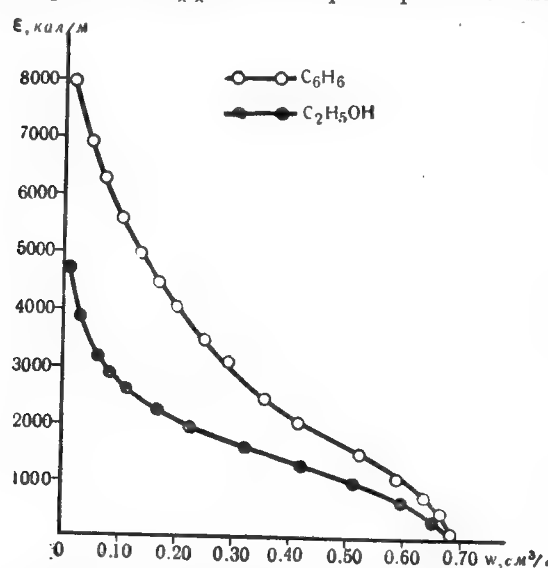
Фиг. 3. Характеристическая кривая для адсорбции паров бензола на активном угле

На фиг. 3 приведена характеристическая кривая для адсорбции бензола на активном угле, вычисленная по изотермам адсорбции для различных температур [10]. Температурная инвариантность характеристической кривой позволяет вычислять изотермы адсорбции пара для разных температур по изотерме адсорбции пара для одной температуры. Результаты вычислений находятся в хорошем соответствии с опытными данными (см., например, [10]).

Адсорбционные потенциалы для различных парообразных веществ при одинаковых объемах w находятся в постоянных отношениях. Так, для двух паров при $w_2 = w_1$

$$\frac{\varepsilon_2}{\varepsilon_1} = \beta, \qquad (11)$$

где β — коэффициент аффинности характеристических кривых. На фиг. 4 изображены аффинные характеристические кривые для бензола и этилового спирта [10]. В данном случае $\beta = 1.64$. Для активных углей с нормально развитой пористостью коэффициенты аффинности не зависят от образца угля.



Фиг. 4. Аффинные характеристические кривые для адсорбции паров бензола и этилового спирта на активном угле

Изложенные представления о природе адсорбции парообразных веществ в основном были развиты Поляни [6].

Для активных углей с очень мелкими микропорами, например типа угля 1 (табл. 2), имеющая значение для адсорбции поверхность зависит от размеров адсорбируемых молекул. Более мелким молекулам метилового спирта доступны поры угля, в которые не могут проникнуть молекулы бензола. Это сказывается на повышении среднего адсорбционного потенциала для метилового спирта или на относительном возрастании коэффициента аффинности. В табл. 4 приведены отношения адсорбционных потенциалов молекул метилового спирта и бензола β для образцов активных углей с закономерно возрастающими средними размерами микропор при переходе от образца В1 к В6 [8]. Через γ обозначено отношение величин адсорбции гептиловой и пропионовой кислот из водных растворов; чем меньше γ, тем мельче поры угля.

Эти опытные данные непосредственно подтверждают допущение о повышении адсорбционных потенциалов в микропорах угля с размерами порядка молекулярных. Из табл. 4 также следует, что для активных углей с широким интервалом средних размеров микропор коэффициенты аффинности практически не зависят от образца угля.

Продукт обугливания сахара — сахарный кокс, прокаленный без доступа воздуха, не обладает первичной пористой структурой, свойственной другим исходным органическим материалам, например древесине, так как обугливанию сахара предшествует стадия плавления его. Поры кокса очень мелки. Однако путем прогрессирующего активирования кокса углекислым газом можно получить образцы активных углей с все более и более крупными микропорами (ряд углей серии В табл. 4). Об изменении характера пористости образцов можно судить по увеличению значения γ.

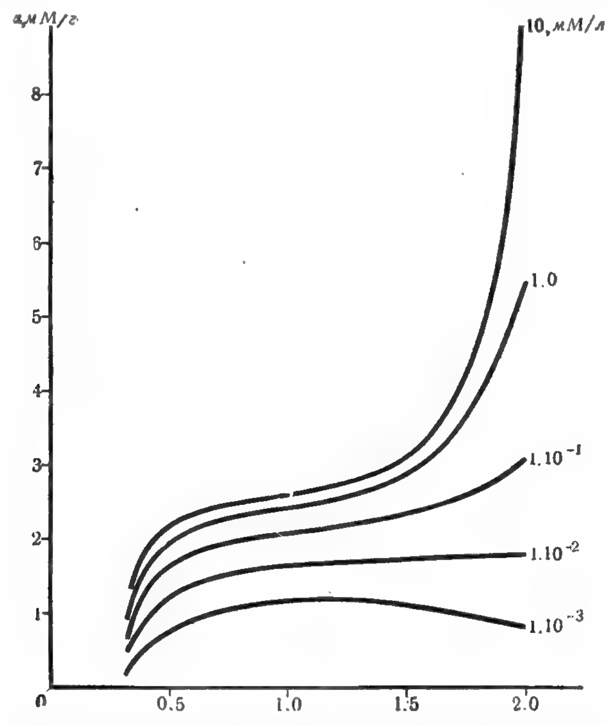
Таблица 4

РАЗМЕРЫ МИКРОПОР И КОЭФФИЦИЕНТЫ АФФИННОСТИ

Уголь	γ	$\beta = \frac{\varepsilon\ CH_3OH}{\varepsilon\ C_6H_6}$
B1	0.33	0.41
B2	0.40	0.37
B3	0.50	0.34
B4	1.6	0.34
B6	2.0	0.33

Влияние характера пористости активных углей на адсорбцию паров бензола иллюстрируется на фиг. 5 кривыми равных равновесных концентраций (изопикнами), выражающими зависимость величин адсорбции от характеристики пористости γ. По мере прогрессирующего активирования, т. е. увеличения γ, возрастает удельная поверхность образцов активных углей. Однако только одно увеличение поверхности угля без изменения характера пористости привело бы к пропорциональному возрастанию адсорбции. В действительности форма кривых резко различна. Для области малых равновесных концентраций порядка тысячных долей миллимоля на литр наблюдается оптимум адсорбции при некотором развитии пористости. При высоких равновесных концентрациях величины адсорбции непрерывно возрастают и в особенности резко в интервале больших γ, когда поверхность пор становится одинаково доступной для молекул исследуемых кислот и значительное увеличение обгара угля при активировании уже не сопровождается существенным изменением отношения γ.

Адсорбция в области малых концентраций в основном происходит на наиболее активных участках поверхности, т. е. в мелких порах молекулярных размеров, где адсорбционные потенциалы повышены. Вначале общее количество таких пор возрастает. При дальнейшем увеличении средних размеров пор эффект повышения адсорбционных потенциалов



Фиг. 5. Изопикны сорбции паров бензола при 20°С активными углями с различным характером пористости

ослабевает и, несмотря на возрастание поверхности угля, величины адсорбции уменьшаются. При высоких концентрациях заполняются относительно крупные поры, и величины адсорбции по существу определяются объемом этих пор, который увеличивается по мере активирования угля. Таким образом, адсорбция парообразных веществ активными углями в значительной степени определяется их структурой.

Структурные особенности каждого образца активного угля могут быть учтены уравнением распределения объема адсорбционного пространства w по адсорбционным потенциалам ε. С этой точки зрения, характеристическая кривая (10) имеет смысл статистической кривой [11]. Принимая гауссово распределение, мы получим уравнение (10) в виде

$$w = w_0 e^{-k\varepsilon^2}, \tag{12}$$

где w_0 — объем ожиженного пара при $\varepsilon = 0$ или при состоянии насыщения адсорбента $\left(\frac{p}{p_s} = 1\right)$. Согласно (12), уравнение характеристической кривой примет вид

$$\varepsilon = \varkappa \sqrt{\ln \frac{w_0}{w}}, \tag{13}$$

где $\varkappa = k^{-\frac{1}{2}}$.

Из двух констант $\varkappa$ и w_0 уравнения (13) только первая зависит от природы адсорбированного пара, так как объем w_0 определяется только структурой адсорбента (w_0 — сумма объемов микро- и конденсирующих, или переходных, пор угля) и, по правилу Гурвича, он в удовлетворительном приближении постоянен для различных паров. Таким образом, из уравнения (13) непосредственно вытекает следствие об аффинности характеристических кривых, ибо при одинаковых объемах w для двух паров

$$\frac{\varepsilon_2}{\varepsilon_1} = \frac{\varkappa_2}{\varkappa_1} = \beta. \tag{14}$$

Уравнение (13) подтверждается на опыте в очень широком интервале изменения w и ε (для w в несколько десятков раз), причем отклонения $\varkappa$ не превышают 10% от среднего значения [11].

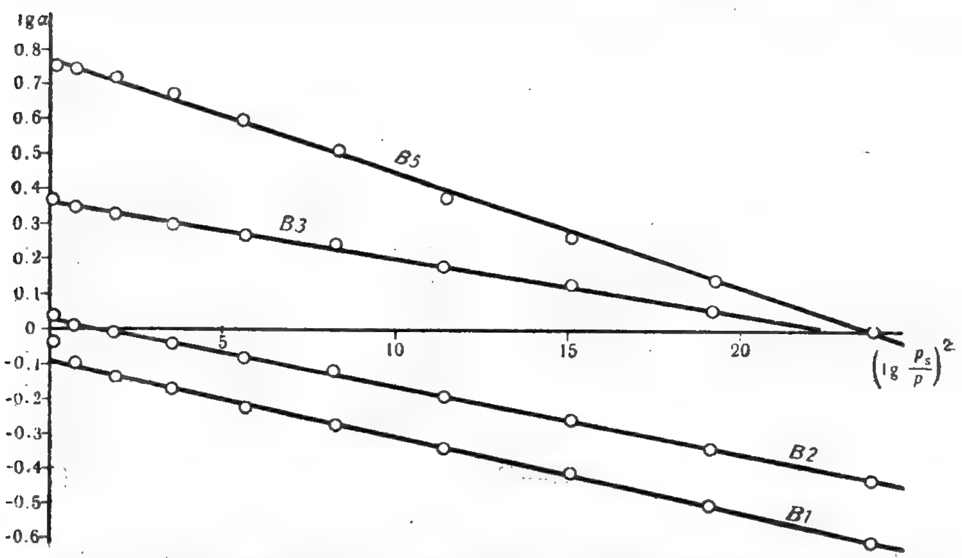
Выразив в независимом от температуры характеристическом уравнении (12) объем w через величину адсорбции a по (7) и адсорбционный потенциал ε — через равновесное давление p по (8), мы получим в связи с (14) $\left(\frac{\varepsilon}{\beta}\text{ не зависит от природы пара для данного адсорбента}\right)$ уравнение изотермы адсорбции для любого пара

$$a = \frac{w_0}{v} e^{-B\frac{T^2}{\beta^2}\left(\lg \frac{p_s}{p}\right)^2} \tag{15}$$

Это уравнение позволяет вычислять величину адсорбции паров при различных температурах, если пользоваться табличными значениями давлений насыщенных паров для каждой данной температуры. Логарифмируя уравнение (15), получаем линейную зависимость между $\lg a$ и $\left(\lg \frac{p_s}{p}\right)^2$

$$\lg a = C - D\left(\lg \frac{p_s}{p}\right)^2, \qquad (16)$$

где $C = \lg \frac{w_0}{v}$ и $D = 0.43\, B\, \frac{T^2}{\beta^2}$. На фиг. 6 изображены в линейной форме,



Фиг. 6. Изотермы адсорбции паров бензола при 20°C различными активными углями, изображенные в линейной форме по уравнению (16)

согласно (16), изотермы адсорбции паров бензола при 20° C для углей В1, В2, В3 и В5, резко различающихся по характеру пористости (см. табл. 4; для угля В5 $\gamma \sim 2$). Экспериментальные точки вполне удовлетворительно ложатся на прямые, что служит подтверждением не только уравнения (15), но и уравнения характеристической кривой (12). Заметим, что формула (16) соответствует опытным данным в широком интервале равновесных относительных давлений (p/p_s) — от $4 \cdot 10^{-5}$ до 0.4.

Из теории Лондона и Поляни следует, что отношение адсорбционных

потенциалов двух парообразных веществ, или коэффициент аффинности, приблизительно прямо пропорционально отношению поляризуемостей молекул α

$$\beta = \frac{\varepsilon_2}{\varepsilon_1} = \frac{\alpha_2}{\alpha_1} \text{ при } w_2 = w_1. \tag{17}$$

Для взаимодействия молекул вследствие проявления дисперсионных сил имеет значение образование мгновенных диполей, т. е. поляризация обязанная флюктуирующему электрическому полю, которое создают соседние молекулы. В данном случае поляризуемость молекул аддитивно складывается из поляризуемостей атомов. Различные модели молекул приводят к приближенному выводу о пропорциональности поляризуемостей молекул и их объемов или молярных объемов веществ в жидком состоянии. Поэтому, в соответствии с (17), можно предположить, что коэффициент аффинности характеристических кривых может быть выражен отношением молярных объемов v парообразных веществ в жидком состоянии [12]

$$\beta = \frac{\varepsilon_2}{\varepsilon_1} = \frac{v_2}{v_1}. \tag{18}$$

Это следствие достаточно удовлетворительно подтверждается на опыте. В табл. 5 сравниваются экспериментальные значения β с вычисленными по формулам (17) и (18), причем все коэффициенты аффинности найдены или вычислены по отношению к характеристической кривой бензола [12].

Таблица 5

КОЭФФИЦИЕНТЫ АФФИННОСТИ ХАРАКТЕРИСТИЧЕСКИХ КРИВЫХ

Парообразное вещество	β (опыт)	$\frac{\alpha_2}{\alpha_1}$	$\frac{v_2}{v_1}$	Парообразное вещество	β (опыт)	$\frac{\alpha_2}{\alpha_1}$	$\frac{v_2}{v_1}$
C_6H_6	1	1	1	C_2H_5Cl	0.76	0.62	0.80
C_6H_{12}	1.03	0.94	1.21	CH_3OH	0.40	0.26	0.46
C_7H_{16}	1.59	1.31	1.65	C_2H_5OH	0.61	0.48	0.65
$C_6H_5CH_3$	1.25	1.27	1.19	CH_3COCH_3	0.88	0.61	0.82
CH_3Cl	0.56	0.44	0.59	CS_2	0.70	0.85	0.68
$CHCl_3$	0.86	0.73	0.90	NH_3	0.28	0.22	0.30
CCl_4	1.05	1.02	1.09				

37*

Уравнение изотермы адсорбции (15) в связи с приближенным выражением для коэффициента аффинности β (18) позволяет решить задачу о зависимости адсорбируемости парообразных веществ активными углями от физических констант паров. Так как по (18), $\frac{\varepsilon}{v}$ = const, то, в связи с (15), получаем приближенное уравнение изотермы адсорбции для различных паров на активном угле [13]

$$a = \frac{w_0}{v} e^{-k \frac{T^2}{v^2} \left(\lg \frac{p_s}{p} \right)^2}, \qquad (19)$$

где w и k — константы, зависящие только от свойств адсорбента. О соответствии этого уравнения опытным данным можно судить по табл. 6, в которой величины адсорбции вычислены только по табличным значениям констант веществ — молярным объемам v и упругостям насыщенных паров p_s. Коэффициенты $w_0 = 0.407$ и $k = 0.0107$ вычислены по изотерме адсорбции паров бензола. Они соответствуют величинам адсорбции в молях на грамм.

Таблица 6

ВЕЛИЧИНЫ АДСОРБЦИИ (мМ/г) ПРИ 20°С И РАЗЛИЧНЫХ РАВНОВЕСНЫХ ДАВЛЕНИЯХ

Парообразное вещество	$p = 0.1$ Hg		$p = 1.0$ Hg		$p = 10.0$ Hg	
	опыт	вычисл.	опыт	вычисл.	опыт	вычисл.
C_6H_6	0.87	1.82	3.05	3.09	4.20	4.17
C_6H_{12}	1.70	1.98	2.55	2.46	3.30	3.55
$C_6H_5CH_3$	2.52	2.45	3.20	3.31	3.80	3.77
C_7H_{16}	1.99	2.04	2.46	2.50	2.93	2.75
C_2H_5Cl	0.40	0.33	1.10	1.15	2.55	2.82
$CHCl_3$	1.30	1.23	2.57	2.63	3.90	4.17

В табл. 6 приведены результаты опытов и вычислений только для некоторых веществ. Столь же удовлетворительное соответствие получено и для других паров, за исключением метилового спирта, где наблюдаются более значительные отклонения.

Таким образом, адсорбируемость парообразных веществ определяется не одной физической константой, как это пытались установить ранее,

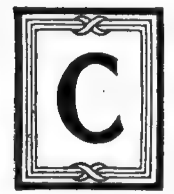
а сочетанием нескольких констант — в рассмотренном приближении молярного объема и упругости насыщенного пара. Адсорбируемость паров зависит также и от структуры активного угля, т. е. распределения объема адсорбционного пространства w по адсорбционным потенциалам ε, выражающимся константами w_0 (сумма объемов микро- и переходных пор) и k. Легко также показать, что отношение адсорбируемости двух паров не есть постоянная величина, она зависит от равновесного давления, при котором сравниваются величины адсорбции. Все это вытекает из уравнения (19).

ЛИТЕРАТУРА

1. *М. М. Дубинин.* Физико-химические основы сорбционной техники. Госхимиздат, изд. 2, 1935.
2. *А. В. Киселев* и *К. Д. Щербакова.* ДАН СССР, т. 45, 1944, стр. 257.
3. *А. В. Киселев, В. П. Древинг* и *А. Д. Рунов.* ДАН СССР, т. 46, 1945, стр. 310.
4. *О. М. Джигит, М. М. Дубинин, А. В. Киселев* и *К. Д. Щербакова.* ДАН СССР, т. 54, 1946, стр. 141.
5. *А. В. Киселев, К. Д. Щербакова.* Acta Physicochimica URSS, vol. 21, 1946, p. 539.
6. *M. Polanyi.* Trans. Farad. Soc., vol. 28, 1932, p. 316.
7. *М. Дубинин.* ЖРХО, т. 62, 1930, стр. 1829; Z. phys. Chem., Bd. 150, 1930, S. 145.
8. *М. Дубинин* и *Е. Заверина.* Acta Physicochimica URSS, vol. 4, 1936, p. 647.
9. *М. Дубинин.* ЖОХ, т. 1, 1931, стр. 289; Z. phys. Chem., Bd. 155, 1931, S. 116.
10. *М. Дубинин* и *К. Чмутов.* Физико-химические основы противогазового дела. М., 1939.
11. *М. Дубинин* и *Л. Радушкевич.* ДАН СССР, т. 55, 1947, стр. 331.
12. *М. Дубинин* и *Д. Тимофеев.* ДАН СССР, т. 54, 1946, стр. 705.
13. *М. Дубинин* и *Д. Тимофеев.* ДАН СССР, т. 55, 1947, стр. 141.

Член-корреспондент АН СССР
С. Н. Данилов

ЗНАЧЕНИЕ РАБОТ А. Е. ФАВОРСКОГО В ИССЛЕДОВАНИИ РЕАКЦИЙ ОДНОВРЕМЕННОГО ВОССТАНОВЛЕНИЯ-ОКИСЛЕНИЯ



Среди многочисленных замечательных работ академика Алексея Евграфовича Фаворского (1860—1945) многие относятся к советскому периоду его научной деятельности. По своей теоретической углубленности, широте охвата химических проблем и практическому значению очень интересны исследования реакций одновременного восстановления и окисления. В своей работе 1928 г. под заглавием «Реакция одновременного восстановления и окисления и изомерные превращения» [1], являющейся центральной работой в только что названной серии его исследований, Фаворский подразделил реакции одновременного восстановления и окисления на три группы:

1) изомерные превращения, сопровождающиеся выделением воды (при дегидратации альфа-гликолей и пинаконов, в частности);

2) изомерные превращения, сопровождающиеся присоединением элементов воды (реакция Канниццаро);

3) изомерные превращения, совершающиеся с сохранением состава, а иногда и функции исходной молекулы; сюда относится подробно описанное в работе 1928 г. изомерное превращение альфа-кетоспиртов (Фаворский, 1914—1937), превращение альдегидов [2] в кетоны и кетонов [3] в изомерные кетоны, а также ряд других аналогичных реакций, не названных Фаворским в этом перечне реакций; эти превращения целесообразно называть реакциями внутримолекулярного окисления-восстановления.

Характерная особенность реакций внутримолекулярного окисления-восстановления заключается в такой перестройке молекулы без изменения состава, при которой происходит перераспределение водородных и кислородных атомов и гидроокислов (если ограничить рассмотрение этих реакций такими веществами, которые состоят только из трех элементов — углерода, водорода и кислорода).

В сочетании с реакцией Канниццаро реакции внутримолекулярного окисления-восстановления весьма распространены при разных химических и биохимических превращениях, в частности при процессах спиртового и других видов брожения сахаров.

При реакциях одновременного окисления и восстановления происходит или внутримолекулярное окисление одного углеродного атома и восстановление другого углерода одной и той же молекулы, как при изомеризации кетоспиртов:

$$CH_3CH_2CO\ CHOH\ CH_3 \rightarrow CH_3CH_2\ CHOH\ CO\ CH_3,$$

или же окислительно-восстановительный процесс разыгрывается с обязательным присоединением воды, причем в реакции участвует две молекулы или одна, как при реакции Канниццаро для бензальдегида и глиоксаля:

$$2C_6H_5C{\overset{O}{\underset{H}{\lessgtr}}} \rightarrow C_6H_5CH_2OH + C_6H_5COOH.$$

Первая указанная Фаворским группа превращений, имеющих место при отнятии воды от гликолей, характеризуется тем, что при этом образуется новая молекула, у которой один углерод более сильно окислился, а другой оказался более восстановленным, чем в исходной молекуле:

$$CH_2OH\ CH_2OH \rightarrow CH_3\ C{\overset{O}{\underset{H}{\lessgtr}}}.$$

Одновременно и независимо друг от друга А. Е. Фаворский и Я. И. Михайленко рассмотрели [4] реакцию дегидратации гликолей как реакцию одновременного окисления и восстановления.

А. Е. Фаворский в разные годы дал очень яркие примеры, в особенности, в отношении третьей группы реакций одновременного окисления-восстановления и, отчасти, первой группы таких реакций (по его клас-

сификации). Ему же принадлежит и схема реакции Канниццаро через промежуточные оксиэфирные группировки.

Область реакций одновременного окисления-восстановления, хотя и рассмотренная в то время Фаворским в несколько иной форме, была плодотворно разработана еще в его докторской диссертации [6]. Как указано в предисловии к диссертации, целью исследования было изучение реакции присоединения хлорноватистой кислоты к двузамещенным ацетиленовым углеводородам для того, чтобы найти характерные для них реакции. Оказалось, что хлорноватистая кислота присоединяется по тройной связи с образованием несимметрических дихлоркетонов. Обрабатывая полученные дихлоркетоны 10% раствором поташа для получения дикетонов или продуктов их циклизации, Фаворский открыл совершавшуюся при этом реакцию синтеза кислот акрилового ряда. Эта реакция была распространена Фаворским на ряд дихлоркетонов из многих двузамещенных ацетиленовых углеводородов, так что были получены различные кислоты этиленового ряда с разветвленными углеродными скелетами (альфа-замещенные акриловые кислоты) и сама акриловая кислота, когда исходным дихлоркетоном явился несимметричный дихлорацетон. Эти превращения можно выразить такими общими формулами:

$$\begin{array}{c} R' \\ | \\ C \\ ||| \\ C \\ | \\ r' \end{array} + 2HClO \rightarrow \begin{array}{c} R' \\ | \\ CO \\ | \\ CCl_2 \\ | \\ r' \end{array} \xrightarrow{H_2O} \begin{array}{c} R'' \\ | \\ C-r' \\ | \\ COOH \end{array} \text{ или } \begin{array}{c} r'' \\ || \\ C-R' \\ | \\ COOH \end{array}$$

где R' и r' — одноатомные жирные радикалы, которые могут быть или одинаковыми, или один из них будет отвечать водороду; R'' и r'' — образующиеся при реакции двухатомные радикалы.

Здесь представлены изомерные превращения, сопровождающиеся перестройкой углеродного скелета и изменением химической функции с присоединением воды и отщеплением хлористого водорода.

Изменение состава молекулы (отщепление хлористого водорода) несколько затемняет характерную для этих превращений реакцию одновременного окисления и восстановления различных частей молекулы дихлоркетона с образованием карбоксильной группы. Сначала Фаворский предположил (1892), что открытые им превращения заключаются в бензильной



584 v 585

перегруппировке, по Либиху, первоначально образующегося дикетона с дальнейшим переходом оксикислоты в кислоту этиленового ряда:

$$\begin{array}{c} CH_3 \\ | \\ CCl_2 \\ | \\ CO \\ | \\ CH_3 \end{array} \rightarrow \begin{array}{c} CH_3 \\ | \\ CO \\ | \\ CO \\ | \\ CH_3 \end{array} \xrightarrow{H_2O} \begin{array}{c} CH_3 \quad CH_3 \\ \diagdown \diagup \\ COH \\ | \\ COOH \end{array} \rightarrow \begin{array}{c} CH_2 \\ \| \\ C - CH_3 \\ | \\ COOH \end{array}$$

Впоследствии, в своей диссертации, А. Е. отказался от этой схемы, так как алифатические дикетоны не превращаются в оксикислоты в условиях реакции Либиха (перегруппировка дикетона бензила в бензиловую кислоту — окси-дифенилуксусную кислоту) и, что главное, при действии поташа на дихлоркетоны получается не одна, а две изомерные кислоты этиленового ряда, чего не должно бы наблюдаться, если бы реакция их образования совершалась через соответствующую дихлоркетону оксикислоту.

А. Е. Фаворский сопоставил и разъяснил ряд реакций, в которых наблюдалось образование карбоксильных групп из галоидозамещенных спиртов и карбонильных соединений, например превращение хлораля в дихлоруксусную кислоту при действии цианистого калия или желтой соли (Валлах, 1877) и при действии окиси серебра (Момене, 1866); но последнее указание Валлах и Фаворский считали сомнительным.

При действии цианистого калия хлораль-гидрат легко превращается, по Валлаху, в дихлоруксусную кислоту. В. Мейер (1877) и Валлах (1877) рассматривали это превращение как результат процесса одновременного окисления и восстановления элементами воды, причем цианистому водороду приписывалась роль катализатора, расщепляющего воду на элементы. Но так как щелочность среды и отщепление хлористого водорода — главнейшие условия в рассматриваемых реакциях, то Фаворский остановился на другом толковании реакции хлораля и аналогичных реакций, рассмотрев их в свете схемы промежуточных реакций, применение которой оказалось чрезвычайно плодотворным и помогло открытию новых реакций, как-то, реакций с ацетонхлороформом и синтеза ацетиленовых спиртов (1905). Он не уделил ни в этих, ни в последующих своих работах того внимания роли водорода и гидроксила молекулы воды в этих реакциях, которого она заслуживает, что стало ясным в наше время, особенно при рассмотрении реакции Канниццаро.

В этой схеме, которую Фаворский предложил для объяснения реакций превращения охлоренных альдегидов и кетонов в кислоты, ценным явилось то, что она наглядно поясняет направление этих превращений и учитывает роль щелочи в этих реакциях. В его схеме роль промежуточных продуктов играют гидратные формы карбонильных соединений и хлорокись.

Для объяснения роли цианистого калия при действии на хлораль и трихлорацетон Фаворский допустил, что цианистый водород промежуточно присоединяется к карбонильной группе и затем отщепляется при действии воды; отщепление хлористого водорода создает двойную связь.

На опыте было обследовано действие цианистого калия на трихлорацетон, которое оказалось сходным с действием 10% раствора поташа, причем в обоих случаях была получена альфа-хлоракриловая кислота:

$$CH_3COCCl_3 \rightarrow CH_2{=}C\,Cl\,COOH.$$

Дихлоруксусный альдегид при обработке 10% раствором поташа дал монохлоруксусную кислоту, но не с бо́льшим выходом, тогда как при цианистом калии выходы кислоты лучшие:

$$CHCl_2C\begin{matrix} \nearrow O \\ \searrow H \end{matrix} \rightarrow CH_2ClCOOH.$$

На основании сопоставления своих опытов с литературными данными, Фаворский записал все эти превращения одной общей схемой через хлорокиси;

$$\begin{array}{c} R' \\ | \\ CCl_2 \\ | \\ CO \\ | \\ r' \end{array} \rightarrow \begin{array}{c} R' \\ | \\ CCl_2 \\ | \\ C\begin{matrix} \nearrow OH \\ \searrow OH \end{matrix} \\ | \\ r' \end{array} \rightarrow \begin{array}{c} R' \\ | \\ CCl \\ |\;\searrow O \\ C\begin{matrix} \nearrow \\ \searrow OH \end{matrix} \\ | \\ r' \end{array} \rightarrow \begin{array}{c} R' \quad r' \\ \searrow\;\swarrow \\ CCl \\ | \\ COOH \end{array} \rightarrow \begin{array}{c} R' \quad r'' \\ \searrow\;/\!/ \\ C \\ | \\ COOH \end{array}$$

или

$$\begin{array}{c} R'' \quad r' \\ \backslash\!\backslash\;\swarrow \\ C \\ | \\ COOH \end{array}$$

где значения радикалов те же, что и выше.

Для подтверждения того, что хлорокиси могут играть роль при рассматриваемых реакциях, Фаворский посвятил в своей диссертации особую главу действию щелочей на галоидозамещенные спирты и вопросу об

изомерных превращениях галоидозамещенных окисей. Интересно превращение трихлорэтилового спирта (Гарцаролли, 1881) при действии едкой щелочи, причем образуется монохлоруксусная кислота, что, по мнению Фаворского, совершается через хлорокись:

$$\begin{array}{c} CCl_3 \\ | \\ CH_2OH \end{array} \rightarrow \begin{array}{c} CCl_2 \\ | \;\rangle O \\ CH_2 \end{array} \rightarrow \begin{array}{c} CH_2Cl \\ | \\ COCl \end{array} \rightarrow \begin{array}{c} CH_2Cl \\ | \\ COOH \end{array}$$

Доказательство участия хлорокисей в превращениях охлоренных альдегидов и кетонов он видел в образовании акриловых кислот из ацетонхлороформа и из других хлороспиртов. Монохлорпроизводные кетонов (монохлорацетон, монохлорметилэтилкетон) превращаются при действии едкого кали (Фаворский, Мякотин, 1901) в пропионовую и изомасляную кислоты:

$$CH_3COCHClCH_3 \rightarrow CH_3\,\underset{\diagdown O \diagup}{COH—CHCH_3} \rightarrow (CH_3)_2\,CHCOOH\,.$$

В более позднее время [5] Э. Д. Венус-Данилова (1941) описала ранее наблюденные ею (1925) превращения альфа-моногалоидокарбонильных соединений в предельные одноосновные кислоты на примере монобромкетонов, например метил-альфа-бром-изопропилкетона в присутствии спиртовой щелочи или углекислых солей бария и калия, получив триметилуксусную кислоту. Описанное в последние годы (1940) получение триметилуксусной кислоты (Эстон и Гринбери) при действии алкоголятов на метил-альфа-бром-изопропилкетон, как новый метод синтеза трехзамещенных производных уксусной кислоты, есть частный случай кислотного превращения различных моногалоидокетонов, описанного ранее Фаворским и его сотрудниками. Фаворский предсказал превращение моногалогеноциклических кетонов в кислоты с сужением цикла (1895) и осуществил позже эти превращения [6]. Так, хлорциклогексанон образует циклопентанкарбоновую кислоту (1914), а монохлорсуберон — циклогексанкарбоновую кислоту (1918) при действии щелочных агентов (А. Е. Фаворский, В. Н. Божовский):

$$\begin{array}{l} CH_2—CH_2—CHCl \\ | \qquad\qquad\quad | \\ CH_2—CH_2—CO \end{array} \rightarrow \begin{array}{l} CH_2—CH_2—CH \\ | \qquad\qquad\quad \rangle O \\ CH_2—CH_2—C—OH \end{array} \rightarrow \begin{array}{l} CH_2—CH_2 \\ | \qquad\quad \rangle CHCOOH. \\ CH_2—CH_2 \end{array}$$

Эту реакцию одновременного восстановления и окисления, как называл ее Фаворский, он разъяснил через спиртоокисные формы по аналогии с хлорокисными формами, которые были использованы в случае превращения дихлоркетонов в кислоты акрилового ряда. На кислотное превращение альфа-бромальдегидов, как на обязательный случай, исходя из схем Фаворского, было указано Даниловым [7] в его работе о дифенилгликолевом альдегиде. На многих примерах галоидоальдегидов открытой цепи и с полиметиленовыми циклами было показано (Данилов, Венус-Данилова), что кислотное превращение бромальдегидов происходит в присутствии углебариевой соли, в особенности гидроокисей свинца и серебра. Старое указание (Момене) о превращении дихлоруксусного альдегида в монохлоруксусную кислоту в присутствии окиси серебра, видимо, соответствует действительности. Оказалось, что галогензамещенные монозы и дисахариды можно превратить в дезоновые или ортосахариновые кислоты в присутствии гидроокиси свинца [8]. Подлинная сахариновая перегруппировка, известная для моноз, осуществляется (Данилов, Венус-Данилова) для самого альфа-оксиизомасляного альдегида в присутствии гидроокиси свинца (плюмбита свинца):

$$(CH_3)_2COH\,C\begin{smallmatrix}\nearrow O\\ \searrow H\end{smallmatrix} \rightarrow (CH_3)_2CH\cdot C\begin{smallmatrix}\nearrow O\\ \searrow OH\end{smallmatrix}\ .$$

Превращение оксиизомасляного альдегида в изомасляную кислоту [9], интересное в связи с кислотным брожением сахаров, — типичнейший пример реакций внутримолекулярного окисления-восстановления с сохранением первоначального состава молекулы, но с резким изменением ее химической функции.

Рассматривая реакции образования акриловых кислот из дихлоркетонов, Фаворский в своей докторской диссертации не оставил без внимания превращения дикетонов, альдегидокетонов и альдегидов при действии щелочи и снова вернулся к этому вопросу в работе 1928 г. [1].

Превращение ароматических дикетонов в оксикислоты с перестройкой углеродного скелета (бензильная перегруппировка), дикетонов полиметиленового ряда в оксикислоты с сужением цикла (Цинке, 1888), превращения ацетилкарбинола и бензоилкарбинола при действии медного купороса в присутствии щелочи в молочную и миндальную кислоты и др.

Фаворский представил через спиртоокисные формы:

$$C_6H_5COCOC_6H_5 \rightarrow C_6H_5\underset{\diagdown\ \ O\ \ \diagup}{COH—COH}C_6H_5 \rightarrow (C_6H_5)_2\,COHCOOH \text{ и др.}$$

Схемами Фаворского можно записать изомерные превращения [7] дифенилгликолевого альдегида, дополнившие ту цепь превращений бензила, первое звено которой было найдено еще Либихом на реакции бензильной перегруппировки (Данилов):

$$C_6H_5CHOHCOC_6H_5 \rightarrow C_6H_5COCOC_6H_5 \rightarrow (C_6H_5)_2\,COHCOOH$$

$$\uparrow \qquad\qquad \uparrow \qquad\qquad \uparrow$$

$$(C_6H_5)_2\,COHC\!\begin{matrix}\diagup\!\!\!\!\!= O\\ \diagdown H\end{matrix}$$

$$\uparrow$$

$$(C_6H_5)_2\,CHCOOH$$

А. Е. Фаворский записывает (1895, 1928) реакцию Канниццаро аналогично реакции бензильной перегруппировки, именно через гидратные и спиртоэфирные промежуточные формы:

$$\begin{array}{c} C_6H_5 \\ | \\ HC\begin{matrix}\diagup OH\\ \diagdown OH\end{matrix} \\ \\ HC\begin{matrix}\diagup OH\\ \diagdown OH\end{matrix} \\ | \\ C_6H_5 \end{array} \rightarrow \begin{array}{c} C_6H_5 \\ | \\ HCOH \\ \diagdown \\ \ \ \ O \\ \diagup \\ HCOH \\ | \\ C_6H_5 \end{array} \rightarrow \begin{array}{c} C_6H_5 \\ | \\ COOH \\ \hline \\ CH_2OH \\ | \\ C_6H_5 \end{array}$$

От двух частиц гидратированного альдегида отнимается одна частица воды, и образующийся диоксиэфир распадается, подобно многоохлоренным простым эфирам, так что «ангидридный кислород отходит всецело к одному углеродному атому, а атом водорода от этого последнего перемещается к тому углеродному атому, с которым только что порвалась кислородная связь» (Фаворский, 1895).

Для объяснения кислотного превращения циклических монохлоркетонов (Фаворский, В. Н. Божовский, 1914—1918), изомерных превра-

щений кетоспиртов и механизма спиртового брожения (1928) Фаворский широко применил схемы через промежуточные спиртоокисные и оксиэфирные формы.

Повторяя свои прежние схемы и уточняя их, он указывает (1928), что «при альдегидах, глиоксале и хлорале восстановление идет за счет перемещающегося альдегидного водорода, а окисление — за счет кислорода воды; при хлоркетонах и бензиле восстановление сказывается перемещением радикалов; промежуточные спиртоокиси, спиртохлорокиси и спиртоэфиры, образующиеся в присутствии щелочи, превращаются далее в кислоты».

Основной экспериментальный материал статьи 1928 г. [1] касается именно изомерных превращений кетоспиртов и здесь принимаются промежуточные спиртоокиси, аналогичные оксоктенолу А. М. Бутлерова (1878)

$$(CH_3)_3C-\underbrace{COH-C}_{O}(CH_3)_2,$$

которые могут возникнуть и без участия воды вследствие таутомерного превращения

$$\begin{array}{c} R \\ | \\ CO \\ | \\ CHOH \\ | \\ R_1 \end{array} \rightarrow \begin{array}{c} R \\ | \\ COH \\ | \;\; \rangle O \\ HC \\ | \\ R_1 \end{array} \rightarrow \begin{array}{c} R \\ | \\ CHOH \\ | \\ CO \\ | \\ R_1 \end{array}$$

$$\begin{array}{c} R \\ | \\ CO \\ | \\ COH \\ \diagup \;\; \diagdown \\ R_1 \quad R_2 \end{array} \rightarrow \begin{array}{c} R \\ | \\ COH \\ | \;\; \rangle O \\ C \\ \diagup \;\; \diagdown \\ R_1 \quad R_2 \end{array} \rightarrow \begin{array}{c} R \quad R_1 \\ \diagdown \;\; \diagup \\ COH \\ | \\ CO \\ | \\ R_2 \end{array}$$

Первое сообщение об изомерных превращениях кетоспиртов было сделано А. Е. Фаворским в 1914 г., а в 1928 г. описано подробно на ряде примеров. Эта интересная реакция была открыта при изучении бромкетонов. В большой работе о действии галоидных соединений фосфора на кетоны, бромкетоны и кетоспирты (1912), Фаворский описал [6] своеобразное действие пятихлористого фосфора на изобутирон, заключающееся

в том, что вместо обычного замещения карбонильного кислорода на два атома галоида, по Фриделю, было наблюдено замещение водорода у прикарбонильного углерода на атом хлора:

$$(CH_3)_2\,CH\,CO\,CH\,(CH_3)_2 \rightarrow (CH_3)_2\,CClCOCH\,(CH_3)_2.$$

Оказалось, что при пятибромистом фосфоре, как и при действии свободного брома на кетоны, всегда происходит именно замещение на бром водорода в альфа-положении к карбонилу.

Получив из третичнобутилэтилкетона при действии пятихлористого фосфора альфа-хлоркетон, А. Е. Фаворский и В. Н. Васильев превратили его в кетоспирт нагревом с муравьинокалиевой солью, но изомерный ожидавшемуся:

$$(CH_3)_3CCOCHClCH_3 \rightarrow (CH_3)_3\,CCHOHCOCH_3.$$

Из бромизобутирона было получено два изомерных кетоспирта вместо одного (А. Е. Фаворский, А. И. Умнова)

$$(CH_3)_2\,CHCOCCl\,(CH_3)_2 \rightarrow (CH_3)_2\,CH\,\underset{\diagdown O \diagup}{COH-C}\,(CH_3)_2 \rightarrow$$

$$\rightarrow (CH_3)_2\,CH\,COH\,(CH_3)\,COCH_3 \leftarrow (CH_3)_2\,CHCOCOH\,(CH_3)_2.$$

Оказалось, что при нагревании с небольшим количеством серной кислоты один кетоспирт прямо изомеризуется в другой.

Такие же изомерные превращения были найдены и на других кетоспиртах: этилбутирилкарбинол изомеризуется с передвижкой карбонила к концу цепи в пропилпропионилкарбинол, а метилпропионилкарбинол — в этилацетилкарбинол (Фаворский, Венус-Данилова):

$$\begin{matrix} CH_3CH_2\,CH_2CO \\ \quad | \\ CH_3CH_2CHOH \end{matrix} \rightarrow \begin{matrix} CHOH\,CH_2\,CH_2\,CH_3 \\ | \\ COCH_2CH_3 \end{matrix}$$

$$\begin{matrix} CH_3CH_2CO \\ \quad | \\ CH_3CHOH \end{matrix} \rightarrow \begin{matrix} CH_3\,CH_2\,CHOH \\ \quad | \\ CH_3\,CO \end{matrix}$$

Такие же взаимоотношения были обнаружены А. Е. Фаворским (совместно с Е. М. Кочергиной) для метилбензоилкарбинола и фенилацетилкарбинола. В более поздней работе [6] (совместно с Т. И. Темниковой) подробно обследованы (1934) свойства этих кетоспиртов, причем этот случай

изомеризации кетоспиртов был назван авторами кетоанольной или кетокарбинольной таутомерией

$$\begin{array}{ccc} C_6H_5-C{=}O & \leftarrow & C_6H_5-CHOH \\ | & & | \\ CH_3-CHOH & \rightarrow & CH_3-C{=}O. \end{array}$$

Эта пара кетоспиртов еще ранее обратила на себя внимание после работ Нейберга (1921), который сочетал уксусный альдегид (пировиноградную кислоту) с бензальдегидом в условиях спиртового брожения и выделил фенилацетилкарбинол вместо более вероятного по схеме синтеза другого кетоспирта — бензоилметилкарбинола. Прямым опытом в лаборатории Фаворского было показано, что в условиях спиртового брожения метилбензоилкарбинол изомеризуется, как и при действии минеральной кислоты, в фенилацетилкарбинол:

$$CH_3\,CHOHCO\,C_6H_5 \rightarrow CH_3COCHOH\,C_6H_5.$$

Как нашли Данилов и Венус-Данилова [10], этот последний кетоспирт получается в результате изомеризации бензилгликолевого альдегида:

$$C_6H_5\,CH_2\,CHOH\,C\langle{}^{O}_{H} \rightarrow C_6H_5\,CHOHCO\,CH_3.$$

Изобутироин в условиях спиртового брожения (Фаворский, Руднева) восстанавливается [6] в диизопропилэтиленгликоль и расщепляется до изомасляного альдегида. Не было найдено изобутилового спирта, что свидетельствует об отсутствии в этих условиях реакции Канниццаро для изомасляного альдегида, но восстановление изобутироина осуществляется именно по реакции Канниццаро при участии изомасляного альдегида. Применив свои схемы для реакций одновременного окисления-восстановления и литературные данные, А. Е. Фаворский (1928) дал интересную схему спиртового брожения.

Сопряженные реакции одновременного восстановления и окисления особенно характерны для этих превращений (вариант 1928 г.):

1) из глюкозы образуются две частицы глицеринового альдегида; при этом имеет место окислительно-восстановительный процесс, так как один углеродный атом по месту распада восстанавливается, давая первичноспиртовую группу, а другой окисляется с образованием альдегидной группы;

2) отщепляя воду по типу дегидратации гликолей, глицериновый альдегид дает метилглиоксаль;

3) молекула глицеринового альдегида и молекула метилглиоксаля

по реакции Канниццаро при участии молекулы воды образуют глицерин и пировиноградную кислоту;

4) из пировиноградной кислоты получается уксусный альдегид и угольный ангидрид;

5) молекула уксусного альдегида и молекула метилглиоксаля дают, по реакции Канниццаро, этиловый спирт и пировиноградную кислоту.

Превращение глицеринового альдегида в последующих стадиях процесса заключается в превращении в метилглиоксаль, который способствует переходу уксусного альдегида в этиловый спирт, а уксусный альдегид непрерывно получается из пировиноградной кислоты, которая все время образуется за счет реакции уксусного альдегида с метилглиоксалем. Поэтому-то имеем глицерин в количестве всего 2—3%.

Различным фактам, наблюденным при брожении при внесении разных веществ в среду, А. Е. Фаворский дал толкование, пользуясь своими схемами, как-то: образованию уксусной кислоты в щелочной среде, синтезу ацилоинов, восстановительным превращениям (карбонильных соединений, нитросоединений и пр.) в бродящей жидкости при участии метилглиоксаля, который при этом образует пировиноградную кислоту. Фаворский считает, что при спиртовом брожении реакции совершаются при действии оксидоредуктазы.

Несколько позже в своем учебнике органической химии (1934) он изменил свою схему, выбросив метилглиоксаль как промежуточный продукт в соответствии с биохимическими данными; вместо него промежуточным продуктом была принята глицериновая кислота, дающая начало пировиноградной

$$C_6H_{12}O_6 \rightarrow CH_2OHCHOH\,C{\overset{O}{\underset{H}{\lessgtr}}}$$

глюкоза — глицериновый альдегид

глюкоза		глицериновый альдегид		
CH_2OH		CH_2OH		CH_3
$CHOH$		$CHOH$		CO
$HCOH$		$COOH$		$COOH$
$>O$	$\rightarrow$	——	$\rightarrow$	$\downarrow$
$HCOH$		CH_2OH		CH_3
$CHOH$		$CHOH$		$C{\overset{O}{\underset{H}{\lessgtr}}} + CO_2$
CH_2OH		CH_2OH		

$$\begin{array}{c} CH_3 \\ | \\ C{\overset{\displaystyle O}{\diagup\!\!\diagup}}\diagdown H \\ \\ C{\overset{\displaystyle O}{\diagup\!\!\diagup}}\diagdown H \\ | \\ CHOH \\ | \\ CH_2OH \end{array} \rightarrow \begin{array}{c} CH_3 \\ | \\ H\,COH \\ \;\;\diagdown O \\ \;\;\diagup \\ H\,COH \\ | \\ CHOH \\ | \\ CH_2OH \end{array} \rightarrow \begin{array}{c} CH_3 \\ | \\ CH_2OH \\ \hline \\ COOH \\ | \\ CHOH \\ | \\ CH_2OH \end{array} \rightarrow \begin{array}{c} CH_3 \\ | \\ CO \\ | \\ COOH \end{array}$$

Таким образом, сложные превращения, совершающиеся при спиртовом брожении, А. Е. Фаворский рассмотрел в свете своих старых схем с участием спиртоокисных и спиртоэфирных форм.

Оксикетонное и кислотное превращения для более сложных оксиальдегидов и оксикетонов — для моноз — были открыты Лобри-де-Брюином в 1895 г. и объяснены через спиртоокисные (альфа-ангидридные формы), что весьма напоминает схемы Фаворского. Оба автора предложили свои схемы независимо друг от друга и почти в одно время (Фаворский, 1894). Но тогда как Фаворский широко разработал свою схему, Лобри-де-Брюин применил свою схему только для случая эпимерной и сахариновой перегруппировки моноз.

Оксикетонное и кислотное превращения оксикетонов (Фаворский) и оксиальдегидов (Данилов, Венус-Данилова) интересны как особенно яркие примеры реакций одновременного окисления-восстановления, совершающегося в одной и той же молекуле без изменения состава. Эти случаи превращений, по классификации Данилова [11], следует относить к реакциям внутримолекулярного окисления-восстановления в противоположность междумолекулярным окислительно-восстановительным превращениям, совершающимся между различными молекулами, например, при участии воды. Бензильная перегруппировка, реакция Канниццаро, процессы гидролитического расщепления карбонильных соединений, ацеталеобразных и сложных эфирных построек и т. п. являются примерами гидролитического окисления-восстановления, сопровождающегося присоединением воды.

Дегидратацию гликолей удобно рассматривать как пример дегидратационного окисления-восстановления, так как процесс сопровождается отщеплением воды. Ярким примером реакций одновременного окисления

и восстановления (таутомерного окисления-восстановления) можно считать таутомерию энолов и карбонильных форм.

Для всех этих случаев одновременного окисления-восстановления можно найти хорошие примеры в исследованиях Фаворского и его школы. Среди примеров одновременного окисления и восстановления, относящихся к изомерным превращениям с сохранением состава (внутримолекулярное окисление-восстановление), Фаворский называет изомерные превращения карбонильных соединений (альдегидов и кетонов).

На примере трифенилуксусного альдегида (1918—1920) С. Н. Данилов обнаружил кетонное превращение альдегидов и поставил эту изомеризацию в связь с ходом дегидратации альфа-гликолей в направлении образования альдегидов и кетонов, а также с ходом изомерных превращений альфа-окисей в альдегиды или кетоны в зависимости от условий (Данилов, Венус-Данилова, 1927). Взаимоотношения карбонильных соединений при отнятии воды от гликолей в зависимости от условий дегидратирования были наглядно представлены на превращениях нескольких альфа-гликолей, как-то: фенилгидробензоин (1919), циклогексилгидробензоин (1925), несимм. метилфенилэтиленгликоль (1927) и др.

Но тогда как альдегиды уже в условиях дегидратации гликолей могут изомеризоваться в кетоны, сами образующиеся кетоны (два кетона из циклогексил-гидробензоина; Данилов, 1925) не изомеризуются один в другой даже в наиболее жестких условиях дегидратирования гликолей. Однако, как впервые доказали А. Е. Фаворский и А. А. Чилингарьян (1926), фенилизопропилкетон при нагревании до 350° с хлористым цинком превращается в несимм. метилфенилацетон, полученный ранее французским химиком Тиффено (1921) из гемдиметилфенилэтиленгликоля

$$(CH_3)_2C(OH){-}CH(OH){-}C_6H_5 \rightarrow (C_6H_5)(CH_3)CH{-}CO{-}CH_3 \rightarrow (CH_3)_2CH{-}CO{-}C_6H_5$$

Позже Фаворский с сотрудниками [6] распространил (1935) реакцию изомерных превращений кетонов на фенилтретичнобутил- и фенилтретичноамилкетоны (над хлористым цинком при 350°).

Рассматривая реакцию дегидратации гликолей как реакцию одно-

38*

временного восстановления и окисления, Фаворский принимал (1928), что направление реакции зависит от большей склонности к окислению одного из двух гидроксилированных углеродов. Основной причиной, направляющей перемещение радикалов при превращении, и является способность к бо́льшему или меньшему окислению одного из двух гидроксилсодержащих углеродных атомов.

По вопросу о путях прохождения реакции дегидратации альфа-гликолей Фаворский высказал плодотворную мысль, что в этих превращениях большую роль играют оксониевые соединения с минеральными кислотами гликолей, а также возможных продуктов дегидратации (1913). Эти соображения были основаны на исследованиях его лаборатории по оксониевым соединениям спиртов и гликолей (о механизме реакций между спиртами и минеральными кислотами, 1913) и на интереснейших свойствах диэтиленового эфира (диоксана), для которого Фаворский еще в 1906 г. нашел удобный способ получения, нашедший себе применение в промышленности (США). Для диоксана им были получены замечательные оксониевые соединения

$$\begin{array}{l} CH_2OH \\ | \\ CH_2OH \end{array} \rightarrow \begin{array}{l} CH_2-O-CH_2 \\ | \qquad\quad\; | \\ CH_2-O-CH_2 \end{array} \qquad \begin{array}{l} \quad\; I \;\; I \\ \quad\;\, \diagdown\!\diagup \\ CH_2-O-CH_2 \\ | \qquad\quad\; | \\ CH_2-O-CH_2 \end{array} \text{ и пр.}$$

$$R''\begin{array}{l} \diagup OH \\ \diagdown OH \end{array} + HX \rightarrow R''\begin{array}{l} \diagup \overset{\displaystyle H}{\overset{|}{O}}-H \;\; \\ \quad\;\; \underset{X}{|} \\ \diagdown OH \end{array} \text{ или } R''\begin{array}{l} \diagup \overset{\displaystyle H}{\overset{|}{O}}-H \\ \quad\; | \\ \diagdown \underset{\displaystyle X}{\underset{|}{O}}-H \end{array}$$

В работах Фаворского за советский период его деятельности можно указать и на другие замечательные теории и факты, которые явились богатейшим вкладом в органическую химию и которые можно рассматривать в группе реакций одновременно и окисления и восстановления.

Продолжая изучение производных ацетиленового ряда, А. Е. Фаворский, признанный мировой авторитет в химии ацетиленовых и диэтиленовых углеводородов, нашел за последние годы интересные пути перехода от ацетиленовых углеводородов к терпенам и дал интереснейшую схему образования терпенов в природе. Эти сложнейшие реакции сопровождаются окислительно-восстановительными превращениями.

Общеизвестно, что еще в прошлом веке и в самом начале текущего столетия А. Е. Фаворский приобрел большую популярность открытыми им реакциями в области ацетиленовых и алленовых углеводородов, а также ненасыщенных спиртов.

Если рассматривать реакции одновременного окисления-восстановления широко, то известные перегруппировки Фаворского однозамещенных ацетиленовых углеводородов в двузамещенные и обратно через алленовые углеводороды можно также причислить к группе окислительно-восстановительных превращений в том смысле, что в молекуле перемещаются центр ненасыщенности (кратные связи) и атомы водорода (в бета-положении).

Реакция взаимодействия, по Фаворскому, ацетиленовых углеводородов с кетонами (1902—1905) явилась способом синтеза [6] ацетиленовых спиртов и нашла себе ценные приложения в синтезе терпеновых спиртов (Ружичка) и для получения синтетических каучуков. Этой же реакцией Фаворский воспользовался и для разъяснения образования терпенов в природе.

В области ненасыщенных углеводородов и спиртов Фаворскому принадлежит за последние десятилетия несколько прекрасных исследований, являющихся дальнейшим развитием его работ в этой области в первую половину его научной деятельности, сразу вписавших его имя в историю мировой органической химии. Нельзя не отметить его работ по изучению возможности введения и степени устойчивости ацетиленовой, алленовой и дивинильной группировок в циклах с 5, 6 и 7 членами (1913—1936). Оказалось, что величина кольца имеет существенное значение, что объясняется стереохимией циклических систем. Тесно примыкает к этим работам интересная статья А. Е. Фаворского и Т. А. Фаворской о порядке отщепления галоидоводородов в открытой цепи с точки зрения стереохимической гипотезы (1922). Никак нельзя умолчать о замечательном докладе А. Е. Фаворского [6] на декабрьской сессии Академии Наук СССР в 1936 г. на тему «Изомерные превращения и явления полимеризации в рядах высоконепредельных углеводородов и их производных». В этом докладе представлены в новом свете изомерные превращения ацетиленовых углеводородов (ацетилен-аллен-диеновая перегруппировка) и сообщено об успехах его лабораторий в области синтетических каучуков (изопренового и метилхлоропренового каучука).

Реакция синтеза третичных ацетиленовых спиртов, по Фаворскому, в присутствии едкого кали получила в руках других исследователей видоизменение в том отношении, что вместо едкого кали был взят амид натрия и реакция была распространена на сам ацетилен. В мировую войну 1914—1918 гг. в Германии был предложен способ получения изопрена из ацетона и ацетилена через диметилацетиленилкарбинол, который гидрировался до диметилвинилкарбинола, а последний при отщеплении воды образовывал изопрен. За последние годы в лаборатории Фаворского реакция конденсации ацетона с ацетиленом была осуществлена не с амидом натрия, а с порошковатым едким кали и притом с лучшими выходами диметилацетиленилкарбинола

$$CH_3-CO-CH_3+CH\equiv CH+KOH\rightarrow(CH_3)_2COH\,C\equiv CH.$$

Полученный третичный ацетиленовый спирт был использован для синтетических каучуков двумя способами:

1. От спирта отнималась вода и к полученному изопропенилацетилену присоединялся хлористый водород при действии конц. соляной кислоты в присутствии полухлористой меди и хлористого аммония, по аналогии с присоединением хлористого водорода (по Ньюленду, 1931) к винилацетилену (это дает хлоропрен для совпрена), причем был получен 2-хлоризопрен, полимеризующийся в метилхлоропреновый каучук, дающий хороший клей

$$(CH_3)(CH_3)C(OH)-C\equiv CH\rightarrow (CH_3)(CH_2{=})C-C\equiv CH\rightarrow (CH_3)(CH_2{=})C-CCl=CH_2\rightarrow$$

$\rightarrow$ метилхлоропреновый (хлоризопреновый) каучук.

2. Второй путь близок к германскому способу получения изопренового каучука (по Мерлингу). Диметилацетиленилкарбинол электролитически гидрировался по способу А. Е. Фаворского и А. И. Лебедевой в этиленовый спирт, при дегидратации которого и получается изопрен, далее полимеризующийся в изопреновый каучук

$$(CH_3)_2C(OH){-}C{\equiv}CH \rightarrow (CH_3)_2C(OH){-}CH{=}CH_2 \rightarrow CH_2{=}C(CH_3){-}CH{=}CH_2 \rightarrow \text{изопреновый каучук}$$

Реакция образования ацетиленовых спиртов из карбонильных соединений и ацетиленовых углеводородов может быть отнесена к реакциям окисления и восстановления, хотя с некоторой условностью. Карбонильный углерод при этом восстанавливается, а ацетиленовый углерод несколько окисляется, присоединяя кислородсодержащий радикал.

Аналогично этому и другая реакция Фаворского, которая приобрела большое значение за границей, а теперь стала широко прививаться на практике и у нас (Фаворский, Шостаковский, 1941), именно реакция образования простых виниловых эфиров, также может быть причислена к реакциям одновременного окисления и восстановления.

Нагревая со спиртовой щелочью аллилен (1888) и аллен (1891), Фаворский получил простой эфир энольной формы ацетона — этилизопрениловый эфир

$$CH_3{-}C{\equiv}CH \rightarrow CH_3{-}C(OC_2H_5){=}CH_2 \leftarrow CH_2{=}C{=}CH_2$$

Подобные ненасыщенные эфиры другого строения получались другим путем А. М. Бутлеровым и А. П. Эльтековым (1877).

Нагревая ацетилен со спиртами в присутствии порошковатого едкого кали, Фаворский и Шостаковский получили различные виниловые эфиры, некоторые из которых уже применялись за границей для разных практических надобностей. Отметим, что, воспользовавшись свойством этих эфиров расщепляться на спирт и на уксусный альдегид при действии сильно разведенной серной кислоты, Фаворский предложил винилэтиловый эфир для получения уксусного альдегида взамен способа гидратации ацетилена (по реакции Кучерова с солями ртути, 1883):

$$CH{\equiv}CH + KOH + C_2H_5OH \rightarrow CH_2{=}CHOC_2H_5 \rightarrow CH_3{-}C{\overset{O}{\underset{H}{\lessgtr}}} + C_2H_5OH.$$

В результате присоединения спирта один ацетиленовый углерод окисляется, приобретая алкоксильную группу, а другой восстанавливается, присоединяя водород. Присоединение кислородсодержащих соединений и воды, например к ацетилену, можно рассматривать как реакцию междумолекулярного окисления-восстановления (гидратационное окисление-восстановление).

Этиленовый спирт — диметилвинилкарбинол, как показано Фаворским и Лебедевой (1938), может претерпевать типичное внутримолекулярное окисление-восстановление в изомерный спирт — изокротилкарбинол

$$(CH_3)_2\,COH\,CH{=}CH_2 \rightarrow (CH_3)_2\,C{=}CHCH_2\,OH.$$

Эта изомеризация — тип аллильной перегруппировки — происходит при действии кислоты.

Изокротилкарбинол играет большую роль в синтезе терпеновых спиртов из диметилвинилкарбинола (по Фаворскому и Лебедевой).

Обрабатывая 20% серной кислотой диметилвинилкарбинол, Фаворский и Лебедева (1938) нашли, что при этом наблюдаются реакции уплотнения в терпеновые спирты — линалоол и, видимо, гераниол. Как известно, линалоол (третичный спирт) в результате аллильной перегруппировки может превращаться в первичный спирт — гераниол.

Фаворский предполагает, что терпены образуются в природе из белков за счет возникающего из них при их гидролизе лейцина, а из него — амилового спирта и изовалерьянового альдегида. Изовалерьяновый альдегид превращается через свою таутомерную форму в изокротилкарбинол и далее — в диметилвинилкарбинол и изопрен

$$(CH_3)_2CH^{*}{-}CH_2{-}CHO \rightarrow (CH_3)_2CH{-}CH{=}CHOH \rightarrow (CH_3)_2C{=}CH{-}CH_2OH \rightarrow (CH_3)_2COH{-}CH{=}CH_2 \rightarrow CH_3(CH_2{=})C{-}CH{=}CH_2$$

Диметилвинилкарбинол является ведущим веществом в процессе синтеза, образуя изопрен, а изомеризуясь в изокротилкарбинол, дает под влиянием кислых катализаторов первичные и третичные спирты, состав которых постепенно осложняется на группу C_5H_8, как-то: линалоол, гераниол,

фарнезол и т. д. Диметилвинилкарбинол и изопропилкарбинол можно считать простейшими терпенами, монотерпенами, из которых образуются более сложные. Эта схема образования терпенов в природе представляет большой интерес среди других гипотез. Проблема образования терпенов еще ждет своего окончательного решения.

В последнем своем публичном выступлении, в докладе [12] на сессии Академии Наук в сентябре 1943 г., А. Е. Фаворский представил интересные данные о таутомерии альдегидов и кетонов ацетиленового ряда, обсуждая вопросы валентности углеводородного сродства. Тройная связь способствует энолизации карбонильных форм, происходит легкое перемещение водорода к карбонильной группе — процесс внутримолекулярного окисления-восстановления.

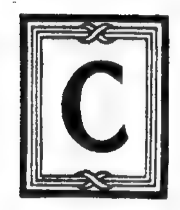
Глубокое изучение многообразных реакций одновременного восстановления и окисления, разработка разделов органической химии о глубоконенасыщенных соединениях, поучительнейшие изомерные превращения, плодотворнейшие синтетические реакции, открытые А. Е. Фаворским, его оригинальные теоретические взгляды — ценнейший вклад в советскую и мировую органическую химию и промышленность.

ЛИТЕРАТУРА

1. *А. Е. Фаворский.* Избранные труды. 1940; ЖРХО, т. 60, 1928, стр. 369.
2. *С. Н. Данилов.* ЖРХО, т. 51, 1919, стр. 97.
3. *А. Е. Фаворский.* ЖОХ, т. 5, 1935, стр. 1728.
4. Труды IV Менделеевского съезда 1925 г. Сообщения о научно-технических работах (А. Е. Фаворский, стр. 165; Я. И. Михайленко, стр. 141).
5. *Э. Д. Венус-Данилова.* ЖОХ, т. 11, 1941, стр. 847; *А. Е. Фаворский, В. Н. Божовский, М. В. Бородулин.* ЖРХО, т. 46, 1914, стр. 618.
6. *А. Е. Фаворский.* Избранные труды. 1940, Изд. АН, «Исследование изомерных превращений в рядах карбонильных соединений, охлоренных спиртов и галоидозамещенных окисей». СПб., 1895, и другие работы.
7. *С. Н. Данилов.* ЖРХО, т. 59, 1927, стр. 1105.
8. *С. Н. Данилов, А. М. Гахокидзе.* ЖОХ, т. 6, 1936, стр. 704.
9. *С. Н. Данилов, Э. Д. Венус-Данилова.* ЖОХ, т. 3, 1933, стр. 559.
10. *С. Н. Данилов, Э. Д. Венус-Данилова.* Ber. chem. Ges., Bd. 63, 1930, стр. 2765.
11. *С. Н. Данилов* и *П. Т. Пастухов.* Реакции гидролитического окисления-восстановления. ЖОХ, т. 17, 1947, стр. 1140.
12. *А. Е. Фаворский.* ЖОХ, т. 14, 1944, стр. 274.

Академик
А. А. Баландин

РАБОТЫ С. В. ЛЕБЕДЕВА В ОБЛАСТИ ОРГАНИЧЕСКОЙ ХИМИИ



ергей Васильевич Лебедев — крупный русский химик органик, деятельность которого развернулась в полном объеме в советское время. Особенно известны его исследования в области полимеризации, ставшие классическими. Значение работ С. В. Лебедева, однако, не ограничивается одной областью полимеризации, но выходит далеко за ее пределы. С. В. Лебедев по справедливости может считаться основоположником новой отрасли технологии — технологии синтетического каучука (СК).

Промышленность СК в крупном масштабе впервые в мире была осуществлена в СССР именно на основе работ С. В. Лебедева и его сотрудников, и СССР вправе гордиться таким своим выдающимся ученым. Синтез каучука, по Лебедеву, основан на каталитических реакциях. Многие другие реакции, изучавшиеся Лебедевым, в том числе гидрогенизация, тоже относятся к числу каталитических реакций. Поэтому С. В. Лебедев оказал значительное влияние также на развитие каталитической химии и в особенности такого важного раздела последней, как органический катализ. В этой области Лебедев продолжал работать до самой своей смерти. Работы его, естественно, сыграли выдающуюся роль и в развитии химии высокомолекулярных соединений, в первую очередь химии каучука.

После смерти С. В. Лебедева промышленность СК развивалась все больше и больше. Были построены новые заводы; опытный завод СК (лит. Б), где были поставлены первые крупные опыты С. В. Лебедева по получению СК из спирта, превратился в большой научно-исследовательский институт синтетического каучука (ВНИИСК), носящий его имя. Был развит и первоначальный процесс, предложенный С. В. Лебедевым. Катализатор значительно усовершенствован; он удовлетворяет теперь самым жестким требованиям промышленности.

Следует отметить, что во время второй мировой войны, когда США были отрезаны от баз натурального каучука, там была создана крупная промышленность СК, мощностью в 1 млн. т в год. Приблизительно для 70% этой продукции сырьем служит этиловый спирт, каталитически перерабатываемый по способу, аналогичному способу Лебедева (для остальных 30% сырьем служат преимущественно бутан и бутилен из нефти).

Претворились в жизнь также идеи С. В. Лебедева о возможности получения каучука с различными свойствами в зависимости от химической природы полимеризуемого вещества. При этом очень важным оказалась возможность получать кополимеры, куда относятся, например, так называемые каучуки буна-S — продукт совместной полимеризации бутадиена и стирола, и буна-N — продукт совместной полимеризации бутадиена и акрилнитрила. Способность к кополимеризации была распространена также на смеси с олефинами, особенно с изобутиленом, откуда возникла целая гамма искусственных материалов — оппанол, бутилкаучук, америпол, вистанекс и т. д., нашедших разнообразное применение. Наряду с этим появились искусственные материалы, стоящие на границе между каучуком и пластмассами. Интересны также кремнийсодержащие каучуки, или силиконы. Естественно, что каталитические и кинетические исследования С. В. Лебедева, произведенные в очень актуальных областях, также получили дальнейшее развитие.

В 20-х годах перед нашей страной встал вопрос о замене импортного натурального каучука отечественным. Каучук—незаменимый материал для автотранспорта, авиации, электротехнической промышленности, не говоря уже о разнообразнейших и распространеннейших предметах широкого потребления, начиная от галош и кончая детскими мячами. Выходом из положения, помимо создания плантаций каучуконосов, явился промышленный синтез каучука. В 1926 г. Высшим советом народного хозяйства был

объявлен всемирный конкурс на получение синтетического каучука. Участники конкурса должны были, помимо описания способа получения СК, представить 2 кг синтетического каучука, изготовленного по этому способу. С. В. Лебедев решил принять участие в этом конкурсе, возглавив группу своих учеников в Военно-медицинской академии в Ленинграде. Имевшиеся в их распоряжении средства были очень ограничены, работа велась во внеурочное время. В качестве исходного недефицитного сырья С. В. Лебедев остановился на этиловом спирте, который в большом количестве можно было получить из картофеля. Еще до Лебедева были попытки получения бутадиена из спирта в смеси с альдегидом над окисью алюминия (Остромысленский); однако выходы получались очень малыми (4—5%). С. В. Лебедев предложил совершенно новый и оригинальный способ, по которому исходным веществом служит только спирт. Он предложил также новый, смешанный катализатор, при проведении паров спирта над которым получался бутадиен с выходом до 25% (по весу).

Суммарная реакция выражается уравнением

$$2C_2H_5OH = CH_2{=}CH-CH{=}CH_2+2H_2O+H_2,$$

Здесь при повышенной температуре происходит дегидрогенизация и дегидратация спирта, а также конденсация двух его молекул в молекулу бутадиена-1,3. Последний представляет собой легко сжижаемый газ.

Вторая стадия получения СК — полимеризация бутадиена — была осуществлена С. В. Лебедевым тоже каталитическим путем; катализатором служил металлический натрий.

К 1 января 1928 г., сроку конкурса, С. В. Лебедевым и его сотрудниками было представлено в запечатанном конверте под девизом «Диолефин» описание нового способа получения СК и требуемый образец в 2 кг. Этот способ, единственный из всех представленных, получил премию. Лаборатории С. В. Лебедева была оказана материальная поддержка со стороны «Резинотреста». С. В. Лебедев обязался в течение двух лет дать техно-экономические показатели для проектирования крупных заводов. Способ Лебедева был проверен сначала на укрупненной установке в лаборатории Ленинградского университета. Установка была выполнена в металле и имела мощность в 1 кг СК в сутки. Последующая проверка происходила на установке одного из московских заводов, а затем на специаль-

ном опытном заводе в Ленинграде. После этого в необычайно сжатые сроки был спроектирован, построен и пущен ряд заводов по этому способу.

Главными трудностями, кроме новизны дела, с которыми приходилось бороться на первых порах, были: большое число (помимо бутадиена) побочных продуктов, образующихся из спирта (см. ниже); конкуренция со способом Бызова, по которому получался бутадиен пиролизом керосина под уменьшенным давлением; сопротивление технологов-резинщиков, недоверчиво относившихся к непривычным рецептам получения резины из нового каучука. Однако благодаря энергии, настойчивости и знаниям С. В. Лебедева, поддержке со стороны партии и правительства, а также упорному труду всего коллектива работников СК новый метод был успешно внедрен в промышленность СССР.

В дальнейшем С. В. Лебедев и его сотрудники очень много работали над реакцией получения бутадиена из спирта. Лебедев создал теорию, объясняющую как эту реакцию, так и многочисленные сопровождающие ее побочные реакции. Согласно этой теории, из этилового спирта на поверхности катализатора в первую очередь образуются ненасыщенные «осколки» $-CH_2-CH_2-$ и $-CH_2-CH(OH)-$, сочетанием и дальнейшим превращением которых и получаются все продукты реакции. Отсюда возник целый ряд работ. Много труда было затрачено на разработку способов утилизации побочных продуктов синтеза бутадиена; при этом было получено много важных для химии результатов. Эта задача в настоящее время решается кардинальным образом — доведением выходов бутадиена почти до теоретических, вследствие чего количество отходов получается очень малым.

Способ, предложенный и осуществленный Лебедевым, породил многочисленные исследования как в лабораториях заводов СК, так и в других научных и исследовательских лабораториях. Он явился также стимулом для проведения других работ по синтезу каучука. В настоящее время мы имеем многообразное разветвление работ, начатых С. В. Лебедевым, в самых разнообразных направлениях.

Успех С. В. Лебедева в области синтеза каучука нельзя рассматривать как случайный; этот успех был подготовлен всей его предыдущей деятельностью в исследовании явлений полимеризации и вообще работой по химии непредельных соединений.

Результаты, полученные С. В. Лебедевым при изучении полимеризации, являются фундаментальными в этой области. Именно С. В. Лебедеву принадлежит мысль о том, что наличие в молекуле системы сопряженных двойных углеродных связей $>C=C-C=C<$ вызывает появление каучукообразующих свойств. С. В. Лебедев впервые установил способность бутадиена-1,3 к термополимеризации. Ему же принадлежит открытие, что синтетический каучук можно получать не только из изопрена и диметилбутадиена, но и из бутадиена. Это открытие имело огромное значение, так как современная промышленность получает синтетические каучуки преимущественно из бутадиена.

С. В. Лебедев установил характер влияния структуры молекул на скорость полимеризации. Эти закономерности сформулированы им в следующих двух положениях:

1. В рядах изомеров при перемещении заместителей от крайних атомов сопряженной системы к средним скорость полимеризации возрастает; при обратном перемещении она убывает.

2. В гомологическом ряду увеличение массы заместителей при средних атомах сопряженной системы повышает, а увеличение при крайних атомах понижает скорость полимеризации, если нагревание вести при соответственных температурах.

Таким образом, углеводород вида

$$H_2C=\underset{\displaystyle R}{\underset{|}{C}}-CH=CH_2$$

полимеризуется быстрее, чем углеводород вида

$$H_2C=CH-CH=CH-R.$$

Если через R обозначить сначала метил, а потом этил, то это в первом случае (верхняя формула) соответствует ускорению, а во втором (нижняя формула) — замедлению полимеризации.

Согласно С. В. Лебедеву, мономер — диолефин с системой сопряженных связей — может превратиться либо в димер, либо в смесь полимеров, содержащих не менее шести исходных молекул; полимеров с промежуточным числом остатков диолефинов не образуется. По Лебедеву, обе эти реакции — параллельные, и их нельзя рассматривать как равно-

весные. Повышение температуры способствует образованию димера вместо полимера. В качестве промежуточного образования для обеих реакций С. В. Лебедев принимает парный комплекс со свободными валентностями, например, в случае бутадиена — свободный радикал

$$-CH_2-\underset{|}{CH}-\underset{|}{CH}-CH_2-CH_2-\underset{|}{CH}-\underset{|}{CH}-CH_2-$$

Этот свободный радикал может дальше либо димеризоваться, либо полимеризоваться.

С. В. Лебедев развил определенные представления о строении веществ, образующихся при димеризации и при полимеризации. При димеризации диолефинов с сопряженной системой связей образуются шестичленные циклы, но не углеводороды с открытыми цепями. При этом если диолефины с сопряженной системой имеют симметричное строение, то может образоваться только один димер, например, в случае бутадиена-1,3 только 1-этенилциклогексен-3. В случае несимметрично построенных диолефинов с сопряженными двойными связями при димеризации возможно образование несколько димерных форм. Так, из изопрена могут образоваться: дипентен, 1,3-диметил-3-этенилциклогексен-6, 1-метил-3-изопропенилциклогексен-6 и 1,4-диметил-4-этенилциклогексен-6. Лебедев действительно обнаружил при термополимеризации изопрена первые два углеводорода, причем второй из них был найден им впервые; третий изомер был позже (в 1924 г.) найден Асканом. Взгляды на димеры изопрена, как на углеводороды открытого строения, которые были высказаны Кондаковым, Гарриесом и Остромысленским, впоследствии были опровергнуты.

Наоборот, при полимеризации, согласно С. В. Лебедеву, образуются длинные цепи. Лебедев еще в 1909 г. экспериментально установил при помощи озонирования, что синтетические каучуки из изопрена, диизопропенила и бутадиена-1,3 состоят из звеньев строения

$$>HC-\underset{\substack{|\\R}}{C}=\underset{\substack{|\\R}}{C}-CH<$$

В этом отношении у синтетических каучуков оказалось полное сходство с натуральным каучуком, для которого подобный факт был установлен Гарриесом в 1904 г.

С. В. Лебедев исследовал также полимеризацию труднодоступных алленовых углеводородов. Алленовые углеводороды полимеризуются по другому типу, отличному от рассмотренного выше. При термической полимеризации алленов образуется непрерывный ряд полимеров — от димера до гексамера.

Для полимеризации алленовых углеводородов С. В. Лебедевым установлены следующие правильности:

1. Скорость полимеризации изомерных углеводородов ряда аллена не зависит от положения замещающих групп.

2. В гомологическом ряду с увеличением массы заместителей скорость полимеризации углеводородов ряда аллена растет, если нагревание вести при соответственных температурах.

Можно видеть, что закономерности при полимеризации алленовых углеводородов действительно оказываются совсем другими, чем при полимеризации углеводородов с системой сопряженных двойных связей.

Согласно Лебедеву, при алленовой полимеризации образуются циклические углеводороды ряда циклобутана:

$$>C=C=C< \quad \text{мономер,}$$

$$\begin{array}{l} >C-C=C< \\ \quad\ | \quad\ \ | \\ >C-C=C< \end{array} \quad \text{димер}$$

и т. д.

Поскольку димер имеет систему сопряженных двойных связей (см. формулу, приведенную выше) и последние образуются также при изомеризации алленовых углеводородов, обладающих кумулированной системой двойных связей, полимеризация алленового типа осложняется полимеризацией дивинилового типа.

Особенное значение имеют исследования С. В. Лебедева по полимеризации этиленовых углеводородов. Эту тему первоначально предложил ему А. Е. Фаворский. После Октябрьской революции С. В. Лебедев вернулся к ней, обогащенный опытом, приобретенным им при изучении полимеризации диолефинов. Сначала, как и в этом последнем случае, С. В. Лебедев применил термополимеризацию, а затем, следуя А. М. Бутлерову, — полимеризацию серной кислотой. Однако вскоре он встал на новый путь, применив в качестве катализатора флоридин (впервые для этой цели предложенный Гурвичем) и проточный метод проведения реак-



С. Лебедев.

608 v 609

ции. Лебедевым было найдено, что полимеризация проходит через димерные, тримерные и т. д. формы вплоть до высокополимеров. Эти формы можно выделить. Таким образом, полимеризация этиленовых углеводородов тоже отличается от полимеризации диолефинов с сопряженными двойными связями.

С. В. Лебедевым и в этом случае установлены закономерности в скоростях полимеризации в зависимости от строения. Исследовав более 20 различных олефинов, он нашел, что (при 20°) над флоридином заметную скорость полимеризации имеют только несимметрично двухзамещенные этилены, а также трехзамещенные этилены. Как сам этилен, так и его однозамещенные и симметрично замещенные полимеризуются очень медленно:

$$CH_2{=}CH_2 \quad \text{медленно}$$
$$R-CH=CH_2 \quad \text{»}$$
$$R^1-CH=CH-R^2 \quad \text{»}$$

$$\begin{matrix} R^1 \\ R^2 \end{matrix} \Big> C{=}CH_2 \quad \text{быстро}$$

$$\begin{matrix} R^1 \\ R^2 \end{matrix} \Big> C{=}CH-R^3 \quad \text{»}$$

Особенно подробно была изучена С. В. Лебедевым полимеризация изобутилена $(CH_3)_2C=CH_2$. Реакция проходит настолько легко, что Лебедев рекомендует ее в качестве лекционного опыта в следующем оформлении. Вертикальная трубка длиной 10 см и диаметром 1.5 см переходит внизу в более узкую часть. В эту трубку на слой стеклянной ваты насыпается флоридин слоем в 2 см. В трубку, сверху вниз, пропускается газообразный изобутилен, который уже при комнатной температуре полимеризуется в жидкость, стекающую в более узкую часть трубки. При реакции происходит разогревание; чтобы его умерить, прибор окружается рубашкой с водяным охлаждением.

Замечательная особенность реакций полимеризации олефинов заключается в их обратимости. При более высокой температуре вблизи (200°) происходит деполимеризация. С. В. Лебедев нашел, что такая деполимеризация проходит очень гладко над силикатами, в частности над флоридином.

Реакции деполимеризации также были подробно изучены С. В. Лебедевым. Эти его работы имеют очень крупное теоретическое и практическое значение. В органической химии мы привыкли, что большинство реакций протекает без того, чтобы был затронут углеродный скелет молекулы. Последний представляется очень прочным образованием. С. В. Лебедевым показано, что уже в сравнительно мягких условиях происходит разрыв углеродной цепи в димерах, тримерах и т. д., где связь между мономерами осуществляется при помощи обычных углерод-углеродных связей. Последние по своей природе не отличаются от подобных связей в мономере.

Работы С. В. Лебедева по деполимеризации также дали значительные практические результаты. Алюмосиликатные катализаторы применяются в каталитическом крекинге нефти, развившемся главным образом в США в грандиозных масштабах. Таким способом получают в настоящее время главную массу высокосортных авиационных бензинов (способ Гудри). Мощность заводов по каталитическому крекингу исчисляется десятками миллионов тонн продукции в год. Экспериментальные исследования С. В. Лебедева, преследовавшие теоретические цели, безусловно сыграли значительную роль в развитии этой отрасли промышленности, где в сложном комплексе одновременно протекающих реакций значительное место принадлежит реакциям деполимеризации, изученным в свое время С. В. Лебедевым.

Из главных результатов, полученных Лебедевым при изучении деполимеризации, можно отметить следующие. Найдено, что тример диизобутилена диссоциирует на димер и на мономер. С повышением температуры равновесие сдвигается в сторону деполимеризации, и потому при более высокой температуре остаются полимерные формы с меньшим молекулярным весом. С ростом молекулярного веса стойкость полимерных форм понижается. Так, при 200° над флоридином в его опытах оставалось 78% диизобутилена, тогда как пентамер в этих условиях уже нацело разлагался; промежуточное положение занимают тример и тетрамер. К сожалению, смерть помешала С. В. Лебедеву закончить эти выдающиеся исследования.

Стремление глубже проникнуть в природу мало устойчивых соединений, какими являются непредельные углеводороды, и всесторонне изучить превращения последних привело Лебедева к исследованию ка-

талитической гидрогенизации. Этот раздел работ также целиком относится к периоду после Октябрьской революции.

Каталитическая гидрогенизация непредельных соединений в жидкой фазе, которая затем приобрела такое большое значение для отверждения жиров, получения высокооктановых бензинов и т. д., была впервые осуществлена в обширных исследованиях нашего соотечественника Фокина (1906—1908). Затем в этой области успешно работали Вильштеттер, Армстронг, Вавон, Ю. С. Залькинд и др. Однако наиболее систематические данные, приведшие к установлению определенных закономерностей, были получены С. В. Лебедевым. Его работы по каталитической гидрогенизации до настоящего времени остаются образцом классического экспериментального исследования в этой области.

С. В. Лебедев проводил гидрогенизацию под атмосферным давлением при взбалтывании в термостате и измерял скорость реакции по объему поглотившегося за полминуты водорода. Катализатором служила платиновая чернь, растворителем для гидрируемого вещества являлся обычно спирт. В этих стандартных условиях было исследовано поведение очень многих непредельных соединений и их смесей.

При гидрогенизации индивидуальных соединений — замещенных этиленов — С. В. Лебедевым были найдены следующие закономерности, касающиеся влияния строения на скорость гидрогенизации. Определяющей здесь явилась степень замещения этилена.

1. Быстрее всех гидрируются однозамещенные этилены, а затем, по мере увеличения степени замещения, скорость гидрирования постепенно падает. Наиболее медленно гидрируются четырехзамещенные этилены.

$R-CH=CH_2$ однозамещенный этилен

$R^1-CH=CH-R^2$, $\begin{matrix}R^1\\R^2\end{matrix}\!\!>C=CH_2$ двухзамещенные этилены

$\begin{matrix}R^1\\R^2\end{matrix}\!\!>C=CH-R^3$ трехзамещенный этилен

$\begin{matrix}R^1\\R^2\end{matrix}\!\!>C=C<\!\!\begin{matrix}R^3\\R^4\end{matrix}$ четырехзамещенный этилен

По Лебедеву, природа заместителей R^1, R^2, ... имеет лишь подчиненное значение и мало сказывается на скорости реакции. Работами Вавона далее было показано, что если заместителями являются цепи изостроения, то скорость гидрирования несколько меньше, чем в случае, если заместители имеют нормальное строение. На усложняющие моменты в случае ароматических заместителей было указано также Б. А. Казанским. Однако, несмотря на исключения, тоже имеющие закономерный характер, рассматриваемое правило Лебедева сохранило до настоящего времени свое значение в качестве руководящего принципа.

2. Если рассматривать кривую скорость гидрогенизации — время (кинетическая кривая), то первый ее участок идет параллельно оси абсцисс, что указывает на первоначальный нулевой порядок реакции. При дальнейшем гидрировании, когда исходного вещества остается мало, кинетическая кривая начинает падать, следуя, как показал еще Фокин, приблизительно первому порядку. С увеличением степени замещения этилена это падение, как нашел Лебедев, начинается ближе к началу гидрирования. Таким образом, для однозамещенного этилена, например, для стирола, кривая имеет высокий горизонтальный участок и крутое падение к концу реакции, а для двухзамещенного, например, для циклогексена, горизонтальный участок кривой расположен ниже, чем в первом случае, и падение к концу реакции идет более полого. Та же тенденция усиливается для трехзамещенного — триметилэтилена. Наконец, для четырехзамещенного — тетраметилэтилена — кривая идет ниже всех остальных, и падение скорости происходит практически уже с самого начала.

Дальнейшие закономерности касаются гидрогенизации смесей. Здесь С. В. Лебедевым было установлено:

3. При гидрогенизации смесей этиленов разных степеней замещения получаются кинетические кривые ступенчатой формы. Сначала гидрируется один компонент смеси, а затем, с другой скоростью, другой компонент. При этом обычно первым гидрируется вещество с меньшей степенью замещения при двойной связи.

4. Этилены одной степени замещения не дают ступенчатых кривых.

Эти теоретически очень важные закономерности оказались в данной области фундаментальными. Сам С. В. Лебедев предложил использовать эти закономерности также в аналитических целях. Для определения степени замещения данного непредельного соединения его следует смешать

с этиленом известной степени замещения и найти из опыта кинетическую кривую гидрогенизации. Если получится ступенчатая кривая, то операцию следует повторить со смесью испытуемого вещества с этиленом другой степени замещения, и так поступать до тех пор, пока не получится гладкая, не ступенчатая кривая. Степень замещения исследуемого вещества совпадает со степенью замещения этого последнего вещества.

Аналогичные, хотя и более сложные закономерности были получены С. В. Лебедевым и его сотрудниками также при исследовании гидрогенизации сопряженных систем, ацетиленовых производных и винилацетилена. Однако эти более сложные случаи были исследованы менее подробно.

Уравнения, описывающие кинетику гидрирования смесей, были выведены лишь значительно позже автором этих строк.

Следует особенно подчеркнуть высокий уровень работ С. В. Лебедева и характерную для них неразрывную связь между теорией и практикой. Научные и технические достижения С. В. Лебедева, наряду с указанными выше особенностями его творчества, всегда будут служить примером для советских ученых.

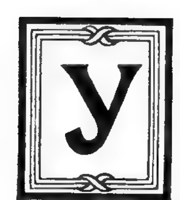
ЛИТЕРАТУРА

Сергей Васильевич Лебедев. Жизнь и труды, Л. 1938.

Академик

Б. А. Казанский

РАБОТЫ АКАДЕМИКА Н. Д. ЗЕЛИНСКОГО И ЕГО ШКОЛЫ В ОБЛАСТИ КАТАЛИТИЧЕСКИХ ПРЕВРАЩЕНИЙ УГЛЕВОДОРОДОВ



Углеводороды содержат в своем составе в различных количественных отношениях всего два элемента — углерод и водород. Это самые простые по качественному составу органические вещества, хотя молекулы некоторых из них могут отличаться очень большой величиной и состоять из многих сотен, а иногда и тысяч атомов углерода и водорода, как это наблюдается, например, в полученных за последнее время высокомолекулярных полимерах этилена и изобутилена, применяемых в качестве пластических масс, изолирующих материалов, пластификаторов и т. д. В простых по качественному составу углеводородах, теоретически говоря, любое число атомов водорода может быть замещено на другие элементы или группы элементов, что дает возможность перехода ко всем остальным известным классам органических соединений, образующим сложную и в то же время единую и стройную систему органической химии, столь пленяющую своим единством и стройностью всякого, кто изучает эту науку или работает в ее области.

Практически такой переход от углеводородов к органическим соединениям других классов осуществляется с различной легкостью для разных групп углеводородов. Если бензол и родственные ему ароматические углеводороды легко вступают в реакции замещения, а этилен и ацетилен вместе с их гомологами могут присоединять к себе хлор, бром, элементы

воды, кислот и некоторых органических соединений и тем самым легко включаются в цепь дальнейших, иногда весьма сложных превращений, то метановые углеводороды — парафины — и циклические — полиметилены, часто называемые циклопарафинами, проявляют большую инертность в химических реакциях и потому они издавна заслужили репутацию «химических мертвецов», не способных к превращению в органические соединения других классов.

Между тем именно эти две последние группы углеводородов составляют основную массу важнейшего углеводородного сырья — естественных горючих газов и нефти, годовая добыча которых во всем мире в последние годы превысила 300 млн. т (кроме СССР). Почти вся продукция нефтяной промышленности, уступающая по своим масштабам только продукции промышленности каменноугольной, используется как топливо или смазка для моторов и двигателей внутреннего сгорания и отчасти как котельное топливо. Только сравнительно небольшая часть ее служит сырьем для химической переработки — для так называемого тяжелого органического синтеза на основе углеводородов.

Таким образом, углеводороды — важнейший источник энергии, значение которого понятно каждому: в нем нуждаются и авиация, и разные виды транспорта, и армия, и военный флот, и сельское хозяйство, и многие отрасли промышленности, и жилища городов. Многие из этих потребителей, особенно авиация и автотранспорт, далеко не безразлично относятся к химическому составу потребляемого ими топлива и смазочных масел. Только некоторые определенные группы углеводородов могут быть достаточно эффективно использованы в современных автомобилях и авиационных моторах. К ним относятся ароматические углеводороды, те парафины, молекулы которых построены из особенно сильно разветвленных цепочек углеродных атомов, некоторые циклические — полиметилены — и непредельные — олефины. Наоборот, в двигателях Дизеля наиболее желательными являются сравнительно высококипящие парафиновые углеводороды нормального строения, обладающие неразветвленными углеродными цепями. Для удовлетворения этих потребностей необходимо вмешательство химии в производство моторного топлива и смазочных масел, и на наших глазах за последние полтора-два десятилетия оно становится все более и более решительным, помогая нефтяной промышленности производить нужные для нее виды углеводородов.

Само собою разумеется, что еще бо́льшую роль играет химия при отыскании путей использования углеводородов как сырья для химической переработки в продукты высшей химической ценности, к которым относятся спирты, уксусная кислота, глицерин, многочисленные галоидопроизводные, простые и сложные эфиры, альдегиды, кетоны, амины, нитросоединения, высшие жирные кислоты, окиси алкенов и т. д. Эти органические вещества имеют первостепенное значение в качестве растворителей для лаков и красок, компонентов при изготовлении пластических масс, исходных материалов для синтеза фармацевтических препаратов, взрывчатых веществ и других не менее важных продуктов.

Ввиду сказанного перед химией углеводородов выдвигаются такие задачи: с одной стороны, найти пути прямого перехода от мало ценных в топливном отношении или трудно поддающихся химическим превращениям углеводородов к углеводородам ценным в топливном отношении или легко используемым как сырье для органического синтеза; с другой стороны, отыскать способы непосредственного превращения таких пассивных в химическом отношении углеводородов, как парафины, в соединения высшей химической ценности.

Вторая задача особенно трудна и решается пока только в единичных случаях. На пути к решению первой задачи сделано гораздо больше, и в настоящее время разные группы углеводородов перестают быть изолированными друг от друга, превращаясь в единый большой класс простейших по качественному составу органических соединений, связанных многими взаимными переходами. От простейшего парафина — метана можно непосредственно перейти к ацетилену, а последний превратить в ароматические углеводороды. От ближайших гомологов метана известен прямой переход к непредельным — олефинам, и далее, к диолефинам, способным полимеризоваться с образованием разных видов синтетического каучука. От парафинов, содержащих в молекуле шесть атомов углерода и больше, можно непосредственно перейти к ароматическим углеводородам. От последних известен простой и легкий переход к гидроароматическим — циклогексану и его гомологам, которые в измененных условиях могут быть превращены обратно в ароматические углеводороды. Циклогексаны и циклопентаны — компоненты любой природной нефти — могут превращаться друг в друга; циклопентаны могут превращаться в парафиновые углеводороды. Почти все перечисленные реакции осуществляются

с помощью каталитических методов и имеют характер непосредственных превращений, не требующих выделения каких-либо промежуточных продуктов.

В разработке многих из этих процессов почетная роль принадлежит русским учёным, в особенности академику Н. Д. Зелинскому и созданной им школе химиков. Дать представление о важнейших из этих реакций является целью настоящего очерка.

Первые работы Н. Д. Зелинского, посвященные каталитическим превращениям углеводородов (в основном каталитической гидрогенизации и дегидрогенизации их), относятся еще к дореволюционному периоду (1911). Эти исследования были прерваны на ряд лет в связи с уходом Н. Д. Зелинского из Московского университета в знак протеста против реакционной политики царского министра Кассо и возобновились только после Октябрьской революции. В советское время они широко развернулись с участием многочисленных учеников Н. Д. Зелинского и дали интересные и важные результаты. На первое место по научному и практическому значению и по богатству экспериментального материала следует поставить работы по каталитической дегидрогенизации циклогексановых углеводородов.

Еще в 1903 г. Сабатье указал, что циклогексан, получаемый присоединением водорода к бензолу в присутствии мелкораздробленного никеля при 160—180°, способен снова переходить в бензол при повышении температуры до 300° и выше; при этом часть бензола превращается в метан. Сабатье дает такое уравнение этой реакции:

$$3C_6H_{12} \rightarrow 2C_6H_6 + 6CH_4.$$

Н. Д. Зелинский заметил, что в присутствии мелкораздробленного палладия и платины дегидрогенизация циклогексана может протекать без глубоко идущей реакции распада с образованием метана и очень точно отвечает уравнению

$$C_6H_{12} \rightarrow C_6H_6 + 3H_2.$$

Так же реагируют метилциклогексан и изомерные диметилциклогексаны.

В присутствии платины и палладия процесс начинается при температуре около 170° и, ускоряясь с ее повышением, при 300° проходит практически нацело. Н. Д. Зелинский и Г. С. Павлов в 1923 г. количественно изучили скорость дегидрогенизации циклогексана при разных темпера-

турах. Позднее Н. Д. Зелинский и А. А. Баландин исследовали кинетику каталитической дегидрогенизации декагидронафталина. Оказалось, что каталитическая дегидрогенизация этих углеводородов протекает очень гладко и подчиняется уравнению Аррениуса, что позволяет вычислить энергию активации этого процесса, составляющую 16 000—19 000 кал/моль для палладиевых и платиновых катализаторов, приготовленных различным образом. Методика изучения кинетики дегидрогенизации гидроароматических углеводородов, разработанная главным образом А. А. Баландиным, отличается большим совершенством, что позволило этой реакции стать одним из наиболее изученных гетерогеннокаталитических процессов в органической химии. Главные выводы упомянутых исследований послужили основой, на которой создалась мультиплетная теория катализа Баландина, являющаяся в настоящее время одной из руководящих теорий в этой области и с течением времени охватившая и объяснившая много новых фактов и явлений.

Начав изучение каталитической дегидрогенизации циклогексановых углеводородов в присутствии палладиевой или платиновой черни, Н. Д. Зелинский постепенно усовершенствовал приготовление дегидрирующих катализаторов, применяя в качестве носителей их асбест, активированный древесный уголь, силикагель. Особенно стойкими и активными оказались катализаторы на активированном древесном угле; они с успехом применялись в многочисленных препаративных и кинетических работах. Никель, в руках Сабатье не давший удовлетворительных результатов, оказался прекрасным катализатором дегидрогенизации циклогексановых углеводородов при нанесении его на гидрат окиси алюминия. Такой катализатор также был изучен с кинетической стороны А. А. Баландиным, А. М. Рубинштейном, Н. И. Шуйкиным и Ю. К. Юрьевым на примерах дегидрогенизации циклогексана, метилциклогексана и диметилциклогексана.

Помимо изучения каталитической дегидрогенизации циклогексана, его гомологов и декалина, Н. Д. Зелинский со своими учениками П. П. Борисовым, И. Н. Тиц, Л. М. Фатеевым и М. В. Гавердовской-Юшкевич наблюдал поведение в этой реакции и более сложных многоядерных углеводородов — гидриндана, дициклогексила, дициклогексилметана, дициклогексилэтана, циклогексилциклопентана. Во всех случаях наблюдалось гладкое превращение циклогексанового кольца в бензольное, иногда

сопровождающееся образованием новых циклов: так из дициклогексилметана получался флуорен

$$C_6H_{11}-CH_2-C_6H_{11} \rightarrow C_{13}H_{10} + 7H_2$$

из дициклогексилэтана — фенантрен

$$C_6H_{11}-CH_2-CH_2-C_6H_{11} \rightarrow C_{14}H_{10} + 8H_2.$$

Каталитическое гидрирование бензола и других углеводородов ароматического ряда в присутствии мелкораздробленного никеля, столь блестяще осуществленное в начале 1900-х годов Сабатье и Сандереном, также привлекло внимание Н. Д. Зелинского. Он показал, что никель вовсе не является катализатором par excellence в этой реакции и что платина и палладий, отложенные на носителях, не только пригодны в этом случае, но имеют ряд преимуществ перед никелем в смысле стойкости, активности, меньшей способности давать побочные продукты и т. д. Таким образом, два класса углеводородов — ароматические и гексагидроароматические — в результате работ Н. Д. Зелинского и его учеников оказались связанными количественно и однозначно гладкими переходами, выражаемыми для простейшего случая бензол-циклогексан уравнением

$$C_6H_6 + 3H_2 \underset{170-300^\circ}{\overset{\text{Ниже } 170^\circ}{\rightleftarrows}} (CH_2)_6,$$

которое показывает, что ниже 170° процесс проходит слева направо, а выше 170° начинает быть заметной реакция дегидрогенизации циклогексана, которая при 300—310° проходит нацело. Присутствие парафинов или циклопентана и его гомологов не влияет заметным образом на течение дегидрогенизации циклогексановых углеводородов, и в этом смысле последний процесс носит избирательный, селективный характер.

Последнее обстоятельство оказалось чрезвычайно важным в практическом отношении, так как оно позволило превращать нацело все циклогексановые углеводороды нефтяных бензинов в ароматические и дало, таким образом, возможность судить о природе циклических полимети-

ленов (нафтенов) в той или иной нефти, поскольку выделить полученные из циклогексанов ароматические углеводороды и разобраться в их природе не составляло особого труда. До этого встречающиеся в нефти полиметилены с пяти- и шестичленными циклами были практически неразделимы, и не существовало никаких способов суждения об их относительном содержании в том или ином бензине. Кроме того, каталитическая дегидрогенизация нефтяных циклогексанов позволила получать из нефти чистые ароматические углеводороды, столь важные в качестве исходных материалов для синтеза красителей, взрывчатых веществ, медикаментов, пластмасс, синтетического каучука и т. д.

Н. Д. Зелинским и его учениками были обследованы многочисленные бензины советских нефтяных месторождений, и в них определено содержание циклогексановых углеводородов, являющихся потенциальным источником ароматических. Из нефтяных «нафтенов», представлявших ранее единую недифференцированную группу, оказалось возможным вполне отчетливо выделить полиметиленовые углеводороды с шестичленными циклами. Число последних в соответствующих фракциях некоторых бензинов оказалось неожиданно большим, что видно из следующих данных Зелинского и Шуйкина:

Фракции	*99.5—102.2°*	*118—120°*	*119.5—121.4°*	*120—124°*	*124—130°*	*130—140°*
Содержание ароматики, об. %: до катализа	ок. 1.0	ок. 1.0	1.5	1.5	3.5	4.5
После катализа	65.0	51.5	54.0	52.0	51.0	43.5
Ароматика, образовавшаяся из циклогексанов, об. %	64.0	50.5	52.5	50.5	47.5	39.0

Данные таблицы представляют собой результаты дегидрогенизации узких фракций сураханского бензина.

Н. Д. Зелинским, а также Н. И. Шуйкиным, Ю. К. Юрьевым, Б. А. Казанским, Г. Д. Гальперном, И. А. Мусаевым и некоторыми другими учениками Н. Д. Зелинского были исследованы методом

каталитической дегидрогенизации бензины (и их отдельные фракции) нефтей балаханской, грозненской, кос-чагылской, калинской, биби-эйбатской, новобогатинской, уральской (Верхне-Чусовских городков); в них было определено содержание циклогексановых углеводородов и обнаружено большое разнообразие в составе бензинов в этом отношении. Б. А. Казанский, Г. Р. Гасан-Заде, Е. И. Марголис, М. И. Маркосова и М. П. Елисеева детально исследовали циклогексановые углеводороды, входящие в состав октанафтеновой и нонанафтеновой фракций калинской и балаханской нефтей с идентификацией отдельных углеводородов.

Реакции каталитического гидрирования бензола и каталитической дегидрогенизации циклогексана своеобразно сочетаются в открытом Н. Д. Зелинским явлении каталитического перераспределения водорода между молекулами циклогексена; одна молекула этого углеводорода дегидрируется, превращаясь в бензол, а две другие, гидрируясь, образуют циклогексан

$$3\ [\text{C}_6\text{H}_{10}] \longrightarrow [\text{C}_6\text{H}_6] + 2\ (CH_2)_6$$

Это превращение, названное Н. Д. Зелинским «необратимым катализом», так как оно проходит только в одном направлении, и бензол с циклогексаном не могут, непосредственно взаимодействуя, образовать циклогексен, протекает в присутствии платины или палладия уже при комнатной температуре.

Замещенные циклогексены также способны реагировать таким образом:

$$3\ [CH_3\text{–C}_6\text{H}_9] \longrightarrow [CH_3\text{–C}_6\text{H}_5] + 2\ CH_3\text{–}CH(CH_2)_5$$

Если двойная связь расположена в боковой цепочке, связанной с циклогексановым кольцом, то, как показали Н. Д. Зелинский и Б. А. Казанский, происходит также реакция необратимого катализа

$$\begin{array}{c} CH_2 \\ \| \\ C \\ CH_2 \quad CH_2 \\ CH_2 \quad CH_2 \\ CH_2 \end{array} \rightarrow \begin{array}{c} CH_3 \\ | \\ C_6H_5 \end{array} + 2 \begin{array}{c} CH_3 \\ | \\ CH \\ CH_2 \quad CH_2 \\ CH_2 \quad CH_2 \\ CH_2 \end{array}$$

Р. Я. Левина, ученица Н. Д. Зелинского, исследовала большое число углеводородов ряда циклогексана, содержащих двойную или даже тройную связь в боковой цепи. В контакте с платиной все они подвергались необратимому катализу, образуя смесь ароматического и сполна гидрированного циклогексанового углеводорода, как видно из следующих превращений аллилциклогексана и циклогексилпропина:

$$3 \begin{array}{c} CH_2 \\ CH_2 \quad CH - CH_2 - CH = CH_2 \\ CH_2 \quad CH_2 \\ CH_2 \end{array} \longrightarrow C_6H_5 - CH_2 - CH_2 - CH_3 +$$

$$+ 2 \begin{array}{c} CH_2 \\ CH_2 \quad CH - CH_2 - CH_2 - CH_3 \\ CH_2 \quad CH_2 \\ CH_2 \end{array}$$

$$3 \begin{array}{c} CH_2 \\ CH_2 \quad CH - CH_2 - C \equiv CH \\ CH_2 \quad CH_2 \\ CH_2 \end{array} \longrightarrow 2\, C_6H_5 - CH_2 - CH_2 - CH_3 +$$

$$+ \begin{array}{c} CH_2 \\ CH_2 \quad CH - CH_2 - CH_2 - CH_3 \\ CH_2 \quad CH_2 \\ CH_2 \end{array}$$

Эти сложные превращения становятся понятными, если допустить, что сначала двойная связь перемещается под влиянием катализатора в кольцо

$$\begin{array}{c} CH_2 \\ CH_2 \quad CH - CH_2 - CH = CH_2 \\ CH_2 \quad CH_2 \\ CH_2 \end{array} \longrightarrow \begin{array}{c} CH \\ CH_2 \quad C - CH_2 - CH_2 - CH_3 \\ CH_2 \quad CH_2 \\ CH_2 \end{array}$$

а затем уже имеет место необратимый катализ образовавшегося замещенного циклогексена. Справедливость такого предположения доказывается тем, что если двойная связь встречает на пути своего перемещения препятствие в виде четвертичного углеродного атома, то необратимый катализ не имеет места, как это наблюдается, например, в случае углеводорода следующего строения:

$$\begin{array}{ccccccc} & CH_2 & & & CH_3 & & \\ CH_2 & & CH & - & \overset{|}{\underset{|}{C}} & - & CH=CH_2 \\ CH_2 & & CH_2 & & CH_3 & & \\ & CH_2 & & & & & \end{array}$$

Если в шестичленном цикле содержатся две двойные связи, то, как показали Н. Д. Зелинский и Г. С. Павлов, необратимый катализ также имеет место, но он протекает в две стадии. Например, из циклогексадиена-1,3 сначала образуются циклогексен и бензол

$$\begin{array}{ccc} & CH_2 & \\ CH & & CH_2 \\ CH & & CH_2 \\ & CH_2 & \end{array} \rightarrow C_6H_6 + \begin{array}{ccc} & CH & \\ CH_2 & & CH \\ CH_2 & & CH_2 \\ & CH_2 & \end{array}$$

а затем циклогексен превращается в бензол и циклогексан по уравнению, приведенному выше.

В течение ряда лет существовало мнение, что дегидрогенизация циклогексановых углеводородов совершенно не затрагивает циклопентанов или парафинов, всегда сопутствующих им в природных бензинах. Предполагалось, что обе последние группы в присутствии катализатора и водорода, выделяющегося из циклогексанов, не способны претерпевать не только дегидрогенизации, но и вообще никаких изменений. Однако в 1934 г. Н. Д. Зелинский, Б. А. Казанский и А. Ф. Платэ показали, что циклопентан при 310° в присутствии платинированного угля гладко превращается в нормальный пентан

$$\begin{array}{ccc} & CH_2 & \\ CH_2 & & CH_2 \\ CH_2 & - & CH_2 \end{array} + H_2 \rightarrow CH_3-CH_2-CH_2-CH_2-CH_3$$

Такому же гидрированию с расщеплением цикла подвергаются разнообразные углеводороды, содержащие циклопентановое кольцо, и, в первую очередь, метил-, этил- и пропилциклопентаны превращаются в изомерные гексаны, гептаны и октаны. В более сложных системах происходит то же. Так С. И. Хромов показал, что α-нафтилциклопентан дает смесь α-амилнафталинов

$$C_{10}H_7-CH\begin{matrix} CH_2-CH_2 \\ | \\ CH_2-CH_2 \end{matrix} + H_2 \rightarrow C_{10}H_7-C_5H_{11}$$

Гидриндан, по наблюдению М. Б. Туровой-Поляк, превращается в орто-этилтолуол

$$C_6H_4\begin{matrix} CH_2 \\ CH_2 \\ CH_2 \end{matrix} + H_2 \rightarrow C_6H_4\begin{matrix} -CH_2-CH_3 \\ -CH_3 \end{matrix}$$

Я. И. Денисенко наблюдал ряд случаев расщепления пятичленного кольца в углеводородах типа

$$C_6H_5-(CH_2)_n-CH\begin{matrix} CH_2-CH_2 \\ | \\ CH_2-CH_2 \end{matrix}$$

$$C_6H_5-(CH_2)_n-CH\begin{matrix} CH_2-CH_2 \\ | \\ CH_2-CH_2 \end{matrix} + H_2 \rightarrow C_6H_5-(CH_2)_n-C_5H_{11}$$

Б. А. Казанский, О. П. Соловова и П. А. Бажулин подробно исследовали гидрирование этил- и пропилциклопентана в присутствии платинированного угля и показали, что расщеплению с присоединением водорода подвергаются все связи пятичленного кольца, но преимущественно — отстоящие через один углеродный атом от углерода, связанного с заместителем. То же было замечено Б. А. Казанским и З. А. Румянцевой для метилциклопентана.

$$\text{R-цикло-}C_5H_9 \;(\text{кольцо: } \underset{|}{\overset{R}{CH}},\ CH_2,\ CH_2,\ CH_2,\ CH_2)$$

$$\xrightarrow{+H_2} R-CH_2-CH_2-CH_2-CH_2-CH_3 \quad \text{(меньше всего)}$$

$$\xrightarrow{+H_2} CH_3-\overset{R}{\overset{|}{C}H}-CH_2-CH_2-CH_3 \quad \text{(преимущественно)}$$

$$\xrightarrow{+H_2} CH_3-CH_2-\overset{R}{\overset{|}{C}H}-CH_2-CH_3$$

Связи, соседние с углеродом, несущим заместитель, разрываются под влиянием водорода в присутствии катализатора в наименьшей степени. Б. А. Казанский, З. А. Румянцева и М. И. Батуев показали, что та же закономерность наблюдается и для двузамещенных циклопентанов. Так, 1,1-, 1,2- и 1,3-диметилциклопентаны реагируют с водородом по следующим схемам:

$$(CH_3)_2C<\text{(кольцо: } CH_2,\ CH_2,\ CH_2,\ CH_2)$$

$$\xrightarrow{+H_2} CH_3-\underset{CH_3}{\underset{|}{\overset{CH_3}{\overset{|}{C}}}}-CH_2-CH_2-CH_3 \quad \text{(связь, соседняя с углеродом, несущим заместитель, совсем не разрывается)}$$

$$\xrightarrow{+H_2} CH_3-CH_2-\underset{CH_3}{\underset{|}{\overset{CH_3}{\overset{|}{C}}}}-CH_2-CH_3$$

$$\text{кольцо: } \underset{|}{\overset{CH_3}{CH}},\ CH-CH_3,\ CH_2,\ CH_2,\ CH_2$$

$$\xrightarrow{+H_2} CH_3-CH_2-CH_2-CH_2-CH_2-CH_2-CH_3$$

$$\xrightarrow{+H_2} CH_3-CH_2-\overset{CH_3}{\overset{|}{C}H}-CH_2-CH_2-CH_3$$

$$\xrightarrow{+H_2} CH_3-\overset{CH_3}{\overset{|}{C}H}-\overset{CH_3}{\overset{|}{C}H}-CH_2-CH_3 \quad \text{(главным образом)}$$

$$CH_3-\underset{|}{\overset{CH_3}{C}}H-CH_2-\overset{CH_3}{\overset{|}{C}}H-CH_3 \xleftarrow{+H_2} \text{(1,3-диметилциклопентан: } CH_3-CH<(CH_2)_3>CH-CH_3 \text{)}$$

(главным образом)

$$\xrightarrow{+H_2} CH_3-\overset{CH_3}{\overset{|}{C}}H-CH_2-CH_2-CH_2-CH_3$$

$$\xrightarrow{+H_2} CH_3-CH_2-\overset{CH_3}{\overset{|}{C}}H-CH_2-CH_2-CH_3$$

Было проведено также сравнение суммарных скоростей присоединения водорода к циклопентану, метилциклопентану и диметилциклопентанам с расщеплением пятичленного кольца. Оказалось, что на скорость этой реакции большое влияние оказывает строение молекулы углеводорода, иначе говоря, число заместителей и их относительное расположение. Быстрее всех гидрируется циклопентан; метилциклопентан — приблизительно в полтора раза медленнее. В 1,1-диметилциклопентане пятичленное кольцо расщепляется примерно с такой же скоростью, как и в метилциклопентане; 1,2-диметилциклопентан реагирует в пять раз, а 1,3-диметилциклопентан — в десять-одиннадцать раз медленнее, чем циклопентан.

Как показали Б. А. Казанский и Т. Ф. Буланова, гидрирование самого циклопентана в присутствии платинированного угля может протекать при гораздо более низкой температуре, чем та, которая указывалась в первых работах по этому вопросу, а именно: уже при 225° в присутствии катализатора средней активности образуется несколько процентов нормального пентана; с повышением температуры его количество быстро возрастает и при 290° и достаточном времени контакта с катализатором может превысить 90%.

Была изучена кинетика реакции и вычислена энергия активации ее, оказавшаяся равной 34 000—35 000 кал/моль.

Таким образом, был найден еще один просто осуществляемый и гладкий переход, связывающий две группы углеводородов, на этот раз — циклопентаны и парафины. Реакция каталитического гидрирования циклопентана при 225° является неожиданной; она не гармонирует с обычным представлением об этом цикле как о весьма устойчивой, свободной от всякого

напряжения системе, очень похожей в этом отношении на парафиновые углеводороды близкого молекулярного веса. Однако ни пентан, ни гексан не подвергаются никаким изменениям в указанных условиях, и потому легкий разрыв связей с присоединением водорода в циклопентане приходится считать явлением, специфичным для углеводородов этой группы, хотя причины его до сих пор не вполне ясны.

В связи со сказанным естественно возникает вопрос, не может ли происходить каталитическое расщепление циклопентановых углеводородов также и при дегидрогенизации циклогексанов в бензинах за счет выделяющегося при этом водорода. Н. Д. Зелинский и Е. М. Шахназарова показали, что при дегидрогенизации циклогексана в смеси с этилциклопентаном из последнего образуется некоторое количество гептанов.

Очевидно, подобного же явления можно ожидать и при дегидрогенизационном катализе бензиновых фракций. Если этот процесс происходит в присутствии достаточно активного платинового или никелевого катализатора, то он должен быть признан селективным только в отношении дегидрогенизации циклогексана; ей сопутствует в большей или меньшей степени расщепление циклопентановых углеводородов. Отыскание такого катализатора, который был бы активен в первой реакции и достаточно пассивен по отношению ко второй, — весьма актуальная задача ближайшего будущего. Такой катализатор оказал бы большую услугу не только для обнаружения и исследования в бензинах циклогексанов путем их превращения в ароматические углеводороды, но он дал бы возможность изучить и циклопентаны, остающиеся при этом неизмененными.

Как указывалось выше, неизменяемость парафиновых углеводородов в присутствии платинового катализатора считалась сама собою разумеющейся. Однако более детальное изучение гидрирования циклопентанов с расщеплением цикла в присутствии платинированного угля привело Б. А. Казанского и А. Ф. Платэ к наблюдению, что при этом образуются бо́льшие или меньшие количества ароматических углеводородов, особенно в тех случаях, когда соответствующий циклопентан имеет боковую цепь из четырех или бо́льшего числа атомов углерода. Исследование природы ароматических углеводородов, образовавшихся из бутилциклопентана, привело к предположению, что они образовались частично путем расши-

40*

рения пятичленного цикла, частично путем циклизации с участием боковой цепи

$$CH_2(CH_2)_3CH\text{—}CH_2\text{—}CH_2\text{—}CH_2\text{—}CH_3 \nearrow (CH_2)_5CH\text{—}CH_2\text{—}CH_2\text{—}CH_3 \rightarrow C_6H_5\text{—}CH_2\text{—}CH_2\text{—}CH_3$$

$$\searrow C_9H_{16} \rightarrow C_9H_{10} \rightarrow CH_3\text{—}C_6H_4\text{—}CH_2\text{—}CH_3$$

Следующим шагом было предположение о возможности непосредственного образования шестичленного цикла из парафинового углеводорода, содержащего не менее шести атомов углерода в цепи. Справедливость этого предположения была подтверждена прямым опытом ароматизации диизобутила в присутствии платинированного угля, в результате чего был получен пара-ксилол

$$CH_3\text{—}CH(CH_3)\text{—}CH_2\text{—}CH_2\text{—}CH(CH_3)\text{—}CH_3 \longrightarrow 4H_2 + CH_3\text{—}C_6H_4\text{—}CH_3$$

Таким образом, был обнаружен еще прямой переход от одной группы углеводородов к другой — от парафинов к ароматическим углеводоро-

дам. Одновременно эта реакция была открыта Б. Л. Молдавским, Г. Д. Камушер и М. П. Кобыльской в ГИВДЕ (Ленинград) и В. И. Каржевым с сотрудниками в ВНИГИ (Москва), но в присутствии катализаторов, содержащих окислы хрома, титана и молибдена и сернистый молибден.

Б. А. Казанский и А. Ф. Платэ нашли, что выходы ароматических углеводородов зависят от строения исходных парафинов. Из нормального гексана было получено в присутствии платинированного угля при 305—310° всего около 2% бензола, а из гораздо более разветвленного диизобутила удалось получить 35% параксилола (при трехкратном контакте с тем же катализатором). Это указывает на то, что наличие боковых цепей в надлежащем положении благоприятствует циклизации. Такой вывод был подтвержден на ряде других примеров.

Ароматизация парафинов в присутствии платинированного угля протекает без образования непредельных углеводородов и других побочных продуктов и может быть изучена кинетически, что было выполнено Б. А. Казанским и А. Л. Либерманом. Они измерили скорость превращения диизобутила в пара-ксилол при разных температурах и нашли, что энергия активации этой реакции составляет 16 000 кал/моль, т. е. очень близка энергии активации дегидрогенизации циклогексана на том же катализаторе. Результатом этой работы был вывод, что ароматизация парафинов в этих условиях протекает через стадию циклогексанового углеводорода

$$\begin{matrix} & CH_2 & \\ CH_2 & & CH_3 \\ CH_2 & & CH_3 \\ & CH_3 & \end{matrix} \xrightarrow{-H_2} \begin{matrix} & CH_2 & \\ CH_2 & & CH_2 \\ CH_2 & & CH_2 \\ & CH_2 & \end{matrix} \xrightarrow{-3H_2} C_6H_6$$

Катализаторы, содержащие окислы металлов, также привлекли внимание лаборатории Н. Д. Зелинского. Б. А. Казанский, А. Ф. Платэ, С. Р. Сергиенко, И. Б. Лосик, А. Л. Либерман, Т. Ф. Буланова и Г. А. Тарасова изучали ароматизацию состоящих главным образом из парафиновых углеводородов фракций синтетического бензина (синтина) в присутствии нескольких десятков различных катализаторов, содержавших

окислы хрома, ванадия, молибдена, тория, урана, никеля, кобальта, при 475—500°. Исследовалась активность и стойкость катализаторов, оптимальные условия проведения процесса ароматизации, свойства получаемых продуктов (содержание ароматических и непредельных углеводородов, октановая характеристика и т. д.). Показано, что синтин, представляющий собою смесь парафиновых и непредельных углеводородов, отличающуюся низкими качествами как топливо для моторов, может быть значительно облагорожен для этой цели путем каталитической ароматизации.

Химизм этого процесса также привлек внимание некоторых учеников Н. Д. Зелинского, особенно А. Ф. Платэ и Г. А. Тарасовой. Ими было показано прежде всего, что этот процесс отличается чрезвычайной сложностью и едва ли может в действительности отвечать тем сравнительно простым схемам, которые обыкновенно для него даются. Повидимому, он слагается из ряда не только последовательно, но и параллельно идущих реакций, некоторые из которых были обнаружены и изучены впервые. К ним относится образование циклогексена из циклогексана на окисных катализаторах и дальнейшая изомеризация его в метилциклопентен, перераспределение водорода в присутствии катализаторов между молекулами циклопентена с образованием циклопентана и циклопентадиена

$$\begin{matrix} CH_2 & — & CH \\ CH_2 & & CH \\ & CH_2 & \end{matrix} \longrightarrow \begin{matrix} CH_2 & — & CH_2 \\ CH_2 & & CH_2 \\ & CH_2 & \end{matrix} + \begin{matrix} CH & — & CH \\ CH & & CH \\ & CH_2 & \end{matrix}$$

весьма напоминающее необратимый катализ циклогексена, образование ароматических углеводородов из гомологов циклопентана и др. Эти исследования в значительной степени расширили наше понимание реакции каталитической ароматизации парафинов.

Взаимные переходы между циклическими углеводородами, содержащими разное число метиленовых групп в циклах, также послужили Зелинскому и его ученикам предметом исследования. Катализатором для этих превращений явился прежде всего безводный хлористый или бромистый алюминий. Цис-форма декагидронафталина при обыкновенной температуре в присутствии хлористого или бромистого алюминия переходит в транс-форму; с повышением температуры, как показали Н. Д. Зе-

линский и М. Б. Турова-Поляк, происходит частичная изомеризация шестичленных колец декагидронафталина в пятичленные. То же, по их наблюдениям, происходит и с циклогексаном

$$\begin{array}{ccc} & CH_2 & \\ CH_2 & & CH_2 \\ CH_2 & & CH_2 \\ & CH_2 & \end{array} \longrightarrow \begin{array}{ccc} CH_2 & - & CH-CH_3 \\ CH_2 & & CH_2 \\ & CH_2 & \end{array}$$

Реакция имеет обратимый характер. Как показал Н. Д. Зелинский еще в 1905 г., этилциклопентан в присутствии бромистого алюминия и избытка брома переходит в метилциклогексан, который далее превращается в пентабромтолуол. М. Б. Турова-Поляк и ряд ее учеников исследовали взаимодействие разнообразных одно- и двузамещенных гомологов циклопентана с хлористым алюминием; оказалось, что все они при температурах не выше 100° в значительной степени, а иногда почти нацело превращаются с расширением цикла в углеводороды ряда циклогексана

$$\begin{array}{ccc} CH_2 & - & CH-CH_2-R \\ CH_2 & & CH_2 \\ & CH_2 & \end{array} \longrightarrow \begin{array}{ccc} & CH_2 & \\ CH_2 & & CH-R \\ CH_2 & & CH_2 \\ & CH_2 & \end{array}$$

Эту реакцию можно применить и к бензинам; содержащиеся в них нафтены с пятичленными кольцами переходят в циклогексаны и тем самым увеличивается потенциальный источник ароматических углеводородов в случае последующей ароматизации по Зелинскому.

Содержащий семичленное кольцо циклогептан, равно как и его гомологи, также переходит при действии хлористого алюминия в углеводсроды ряда циклогексана

$$\begin{array}{l} CH_2-CH_2-CH_2 \\ CH_2-CH_2-CH_2 \end{array} > CH_2 \longrightarrow \begin{array}{ccc} & CH_2 & \\ CH_2 & & CH-CH_3 \\ CH_2 & & CH_2 \\ & CH_2 & \end{array}$$

Подобное же превращение имеет место и для циклооктана

$$\begin{array}{l} CH_2-CH_2-CH_2-CH_2 \\ | \qquad\qquad\qquad\qquad\quad | \\ CH_2-CH_2-CH_2-CH_2 \end{array} \begin{array}{l} \nearrow \\ \\ \searrow \end{array} \begin{array}{c} CH_3 \\ | \\ CH \\ CH_2 \quad CH_2 \\ CH_2 \quad CH-CH_3 \\ CH_2 \\ \\ CH_3 \\ | \\ CH \\ CH_2 \quad CH_2 \\ CH_2 \quad CH_2 \\ CH \\ | \\ CH_3 \end{array}$$

Пожалуй наиболее интересное изменение претерпевает в этих условиях этилциклобутан, испытывающий расширение четырехчленного цикла до пятичленного и далее — до шестичленного

$$\begin{array}{l} CH_2-CH-CH_2-CH_3 \\ | \qquad\quad | \\ CH_2-CH_2 \end{array} \longrightarrow \begin{array}{l} CH_2-CH_2 \\ | \\ CH_2-CH_2 \end{array} \!\!\!> CH-CH_3 \rightarrow \begin{array}{c} CH_2 \\ CH_2 \quad CH_2 \\ CH_2 \quad CH_2 \\ CH_2 \end{array}$$

Таким образом, в присутствии хлористого алюминия при температурах не выше 100° различные циклические углеводороды, содержащие четыре, пять, семь и восемь атомов углерода в кольце, изомеризуются в циклогексан или его гомологи, представляющие собой в этих условиях более устойчивые системы.

Иные отношения наблюдаются между углеводородами с пяти- и шестичленными циклами при более высоких температурах. Н. Д. Зелинский и Ю. А. Арбузов показали, что циклогексен в присутствии окиси алюминия, силикагеля, двуокиси титана, окиси бериллия при 400—450° превращается в метилциклопентен

Метилциклогексен в этих условиях переходит в диметилциклопентен. Значительно более сложный комплекс превращений претерпевает в присутствии тех же катализаторов циклогексадиен-1,3. Н. Д. Зелинский, Ю. А. Арбузов и М. И. Батуев показали, что этот углеводород превращается главным образом в бензол и частично в циклогексен, который далее изомеризуется в метилциклопентен или, вернее, в смесь метилциклопентенов, так как с помощью Раман-спектров теми же авторами было установлено образование из циклогексена трех изомерных метилциклопентенов. Образующийся при реакции водород частично гидрирует циклогексен и метилциклопентены

Циклогептен в присутствии силикагеля изомеризуется в метилциклогексен, который частично дегидрируется в толуол, а частично изомеризуется дальше — в 1,2-диметилциклопентен

$$\begin{matrix} CH_2 - CH_2 - CH_2 \\ | \\ CH_2 - CH_2 - CH_2 \end{matrix} > CH_2 \searrow \text{[cyclo: } CH, CH_2, C-CH_3, CH_2, CH_2, CH_2\text{]} \longrightarrow \begin{matrix} CH_2 - CH - CH_3 \\ | \qquad \searrow C - CH_3 \\ CH_2 - CH \end{matrix}$$

$$\searrow C_6H_5 - CH_3$$

Гептен-1 в присутствии окиси бериллия претерпевает изомеризацию своей прямой углеродной цепочки в разветвленную. Об этом можно заключить на основании того, что после присоединения водорода к продукту изомеризации гептена-1 в нем был констатирован 3-метилгексан

$$CH_2 = CH - CH_2 - CH_2 - CH_2 - CH_2 - CH_3 \rightarrow [CH_3 - CH = \underset{\displaystyle CH_3}{\underset{|}{C}} - CH_2 - CH_2 - CH_3] \rightarrow$$

$$\xrightarrow{+H_2} CH_3 - CH_2 - \underset{\displaystyle CH_3}{\underset{|}{CH}} - CH_2 - CH_2 - CH_3$$

Октен-1 в присутствии окиси бериллия подвергается еще более сильному изменению, и в гидрированном продукте превращения констатированы с помощью Раман-спектров 2-метил-, 3-метил-, и 4-метилгептан

$$CH_2 = CH - CH_2 - CH_2 - CH_2 - CH_2 - CH_2 - CH_3 \longrightarrow$$

$$\longrightarrow \underset{\text{Изооктены}}{[C_8H_{16}]} + H_2 \longrightarrow \begin{cases} CH_3 - \underset{\displaystyle CH_3}{\underset{|}{CH}} - CH_2 - CH_2 - CH_2 - CH_2 - CH_3 \\ CH_3 - CH_2 - \overset{\displaystyle CH_3}{\overset{|}{CH}} - CH_2 - CH_2 - CH_2 - CH_3 \\ CH_3 - CH_2 - CH_2 - \underset{\displaystyle CH_3}{\underset{|}{CH}} - CH_2 - CH_2 - CH_3 \end{cases}$$

Остановимся в заключение на цикле работ, посвященных получению бутадиена путем каталитической дегидрогенизации бутана и бутилена

$$CH_3 - CH_2 - CH_2 - CH_3 \longrightarrow CH_2 = CH - CH_2 - CH_3 + H_2$$
$$CH_2 = CH - CH_2 - CH_3 \longrightarrow CH_2 = CH - CH = CH_2 + H_2$$

и превращению этилбензола в стирол таким же путем

$$C_6H_5-CH_2-CH_3 \longrightarrow H_2 + C_6H_5-CH=CH_2$$

В этих реакциях происходит непосредственное образование двойной связи путем отщепления водорода от двух соседних углеродных атомов в открытой цепи: непосредственный переход от предельных углеводородов к непредельным — олефинам и диолефинам. Он происходит в присутствии катализаторов, состоящих из окислов таких металлов, как хром, ванадий и др., и требует гораздо более высоких температур (500—600°), чем рассмотренная выше дегидрогенизация циклогексанов по Зелинскому.

Бутадиен представляет собою важнейший исходный материал для получения синтетического каучука, и в СССР он в настоящее время добывается из этилового спирта; стирол является важным компонентом при синтезе некоторых видов каучука. Получение бутадиена прямой дегидрогенизацией углеводородов имеет очень важное принципиальное значение, так как связано с возможностью отказа от пищевого сырья. Эта задача привлекла внимание Н. Д. Зелинского и А. А. Баландина. В ряде исследований этого рода были получены прекрасные результаты.

А. А. Баландин, Н. Д. Зелинский, М. Н. Марушкин и Л. Н. Павлов наблюдали, что при 570° в присутствии хромового катализатора бутан приблизительно на 50% превращается в бутилен (более 85% на вошедший в реакцию бутан). А. А. Баландин, Н. Д. Зелинский, О. К. Богданова и А. П. Щеглова показали, что дегидрогенизация бутилена в бутадиен может проходить также в присутствии хромовых катализаторов при температуре около 600°, но только лишь при уменьшенном давлении или при разбавлении углекислым газом или азотом. Бутадиена в таких условиях образуется 30—34% (около 75% на вошедший в реакцию бутилен). Для получения бутадиена можно применять также и получаемую на заводах крекинга нефти техническую смесь бутана и бутилена.

Задача дегидрогенизации этилбензола в стирол также получила разрешение. А. А. Баландин, Н. Д. Зелинский, Г. М. Марукян и О. К. Богданова провели эту реакцию с выходом стирола до 55% в присутствии ванадиевого катализатора при 625° и при разбавлении этилбензола углекислым газом. В дальнейшем А. А. Баландиным и Г. М. Марукян была исследована дегидрогенизация изопропилбензола и многочисленных

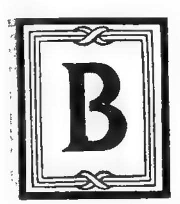
других одно- и двузамещенных гомологов бензола. Было показано, что в большинстве случаев они дегидрируются быстрее и легче этилбензола и получаемые из них замещенные стиролы представляют собой очень ценное сырье для синтеза высокополимерных веществ.

Таковы главнейшие результаты исследований, проведенных Н. Д. Зелинским и его учениками и сотрудниками за последние 30 лет в области каталитических превращений углеводородов. Они во многом изменили наши представления о взаимных отношениях между различными группами этого основного класса органических соединений, объединили и оживили эти группы многочисленными взаимными переходами и наметили новые пути получения некоторых из них. Эта большая работа, потребовавшая много творческих усилий и кропотливого труда, могла быть с успехом выполнена только в связи с тем громадным общим подъемом науки, техники и народного хозяйства в целом, который вызвала в нашей стране Великая Октябрьская социалистическая революция.

Академик

А. А. Баландин

ТЕОРИЯ ОРГАНИЧЕСКОГО КАТАЛИЗА



Введение. Катализ имеет большое значение в современной химии и химической промышленности; достаточно напомнить о каталитическом крекинге, синтетическом бензине, промышленности синтетического каучука, впервые в крупных масштабах осуществленной в СССР. Очень важно поэтому развитие теории катализа. До недавнего времени в объяснении катализа были антагонистами химические теории (промежуточных соединений, деформации молекул) и физические теории (адсорбционные, активных центров). Противоречия между ними устраняются мультиплетной теорией, предложенной автором в 1929 г. Эта теория развивает и конкретизирует воззрения Д. И. Менделеева [59] и Н. Д. Зелинского [53, 9] о катализе как об изменении формы молекул, а также взгляды Лэнгмюра [58]. Мультиплетная теория учитывает тонкую, атомную структуру активных центров, сочетая стереохимические и энергетические соображения. Отсюда — объединение химического и физического подхода к катализу.

Настоящая статья касается мультиплетной теории. Общее развитие органического катализа в СССР рассмотрено в другом месте [55].

1. Принципы мультиплетной теории [1]. Из малого радиуса действия валентных сил мультиплетная теория выводит, что при гетерогенном катализе в молекулах реагируют только атомы, находящиеся в контакте с поверхностью. Для реакции необходимо несколько

поверхностных атомов катализатора — мультиплет. При этом разрыв и образование связей подчиняются определенным правилам. Так, при дегидрогенизации спирта в альдегид с двумя атомами катализатора K соприкасаются, притягиваясь к ним, атомы, которые на фиг. 1 находятся внутри рамки. Атомы молекулы вне рамки направлены от поверхности

а *б* *в*

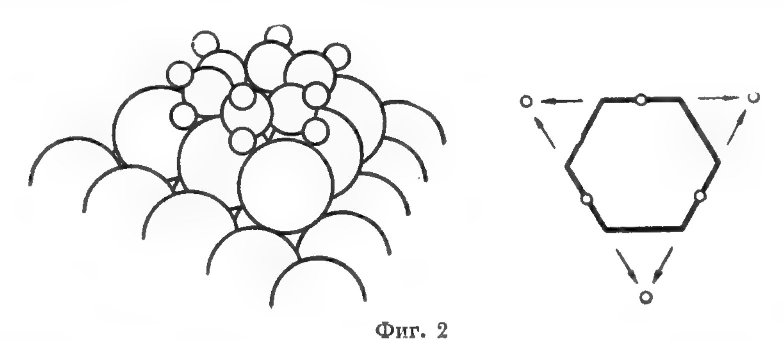
Фиг. 1

катализатора. В результате рассматриваемой дублетной реакции на фиг. 1 две вертикальные связи (*а*) переходят в две горизонтальные (*в*) через нестойкий мультиплетный комплекс (*б*).

Отсюда прежде всего вытекает, что катализируемые молекулы должны находиться в мономолекулярном слое, прилегающем к катализатору; они должны быть ориентированы своими реагирующими частями к катализатору. Такая ориентация молекул спиртов действительно доказана кинетически Констеблом [72] и особенно полно Борком [51]. Специфичность каталитического действия оказывается связанной с разной ориентацией.

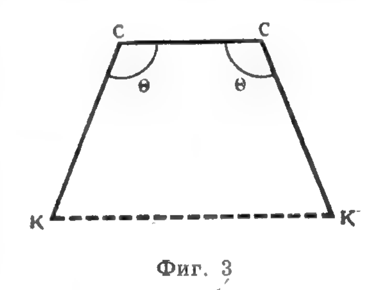
2. П р о с т р а н с т в е н н ы е ф а к т о р ы. Современные знания о форме и размерах молекул и о решетках кристаллов позволяют строить модели мультиплетных комплексов (1). Дегидрогенизация циклогексана и его производных, открытая Н. Д. Зелинским [53, 9], по мультиплетной теории передается секстетной моделью (фиг. 2, внешний вид и в плане). Переходы атомов H указаны стрелками. Имеется соответствие элементов симметрии молекул и катализатора; 6-членное кольцо ложится плоско на грань с треугольным расположением атомов, удовлетворяя принципам, аналогичным указанным выше. Подобный же вид имеет модель обратной реакции — гидрирования бензола.

Модель предвидит, что катализаторами здесь могут быть только металлы и притом только кристаллизующиеся в системах кубической (гранецентрированные кубы, решетка A 1) и в гексагональной (решетка A 3), так как только на гранях (111) первой и (0001) второй имеется требуемое



Фиг. 2

треугольное расположение атомов. Кроме того, атомные радиусы R металлов-катализаторов должны лежать в определенных пределах. Иначе атомы H в циклогексане или будут слишком далеки от притягивающих их атомов катализатора, или кольцо не наложится на решетку. В табл. 1 дан список элементов с указанием их решеток и атомных радиусов [77], причем обведены рамкой элементы, являющиеся катализаторами; можно видеть, что теория полностью подтверждается. Табл. 1 показывает, почему, например, Pt, Pd или Ni являются катализаторами данной реакции, а, например, Ca, Cl или Fe — нет. Недавно найдено [60], что бензол гидрируется на Re, и это подтверждает предвидение теории.



Фиг. 3

Согласно секстетной модели, бензол может гидрироваться также на смешанных кристаллах с возрастающим содержанием одного из металлов, но лишь до тех пор, пока имеется требуемая структура. Это подтверждено рентгенографическими измерениями на смешанных кристаллах Ni, Co, Fe, Pd и Ag [76, 75]. Так, сплав 74.9% Co + 21.7% Fe с решеткой A 1 еще активен, а сплавы 50.0% Co + 50.0% Fe и 24.2% Co + 75.8% Fe с решеткой A 2 уже нет [76].

Таблица 1

АТОМНЫЕ РАДИУСЫ (Å) И РЕШЕТКИ ЭЛЕМЕНТОВ В СВЯЗИ С ИХ СВОЙСТВАМИ КАК КАТАЛИЗАТОРОВ ДЕГИДРОГЕНИЗАЦИИ ЦИКЛОГЕКСАНА

Гранецентриров. кубическая решетка А 1		Объемноцентриров. кубическая решетка А 2		Гексагональная решетка А 3		Решетка алмаза А 4	
Sr	2.248	Cs	2.62	Er	1.865; 1.870	Sn	1.40
Ca	1.966	Rb	2.44	Pr	1.819; 1.828	Ge	1.22
La	1.877	K	2.31	Nb	1.810; 1.828	Si	1.17
Ce	1.818	Ba	2.17	Y	1.797; 1.832	C	0.765
Th	1.795	Na	1.86	Mg	1.595; 1.643	Другие	решетки
Pb	1.746	Zn	1.56	Hf	1.569; 1.600	In	1.62
Tl	1.711	Li	1.52	Cd	1.486; 1.543	Bi	1.55
Ag	1.442	U	1.49			Sc	1.51
Au	1.439	Ta	1.427			Hg	1.50
Al	1.429	Nb	1.426	Re	1.368; 1.377	Sb	1.45
		W	1.367	Zn	1.330; 1.457	Te	1.43
		Mo	1.360	Os	1.335; 1.365	Mn	1.28
		V	1.314	Ru	1.322; 1.350	As	1.25
Pt	1.385	Cr	1.246	Co	1.247; 1.254	Ga	1.22
Pd	1.372	β-Fe	1.239	Ni	1.245; 1.245	Se	1.16
Ir	1.355			Be	1.112; 1.134		
Rh	1.342						
Cu	1.275						
Co	1.254						
Ni	1.244						

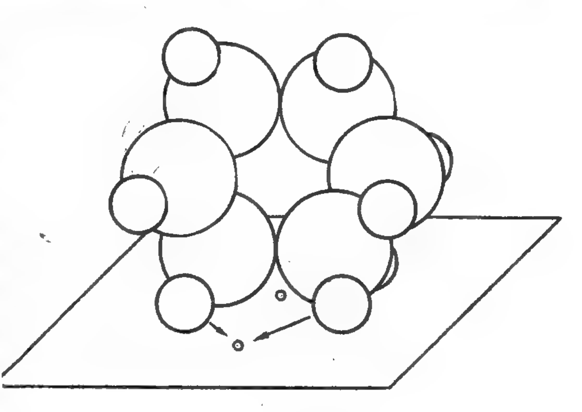
В мультиплетном комплексе валентные углы Θ стремятся сохранить свое нормальное значение, и поэтому и в дублете должно существовать оптимальное для реакции расстояние К—К (фиг. 3) [48].

Железо в виде катализатора, очень активного для синтеза аммиака, не гидрирует бензол, где требуется секстет, но гидрирует этилен, где необходим дублет [75]. Действительно, в решетке железа (А 2) нет секстетов, но есть дублеты с подходящими междуатомными расстояниями.

К дублетному механизму гидрогенизации олефинов (двухточечная адсорбция) приходят также исследователи, изучавшие эту реакцию с применением дейтерия [74].

Для гидрирования этилена Бик [69] установил различную каталитическую активность разных граней никеля, обладающих разными междуатомными расстояниями.

А. М. Рубинштейн [62] рентгенографически и кинетически изучал активность катализаторов из MgO относительно дегидрогенизации и дегидратации бутилового спирта. В зависимости от способа приготовления параметр решетки MgO, т. е. междуатомное расстояние, несколько меняется. Рубинштейном установлено существование оптимального параметра для катализа. Более того, оказалось, что для дегидрогенизации оптимальный параметр меньше, чем для дегидратации, в согласии с тем, что междуатомное расстояние для связи O — H (1.01 Å) меньше, чем для связи C — O (1.46 Å) (ср. табл. 4).



Фиг. 4

Подобно этому, над окислами Nb и Ta ($R = 1.47$ и 1.49 Å) спирты дегидратируются, тогда как над окислами V ($R = 1.36$ Å) преимущественно дегидрируются [21].

Замечена связь между молекулярным объемом окисла и его способностью быть каталитическим промотором [70].

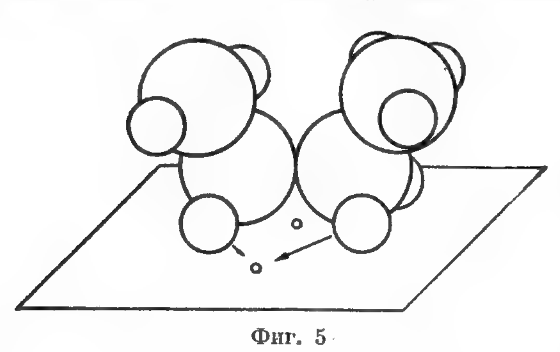
Основываясь на квантово-механической теории активированного комплекса, модели которого очень близки к моделям мультиплетной теории, Эйринг с сотрудниками еще раньше установил [73], что для дублетной адсорбции водорода на Ni и на угле существует оптимальное расстояние Ni — Ni и C — C.

Все эти факты подкрепляют мультиплетную теорию. Ее успехом явилось также нахождение своеобразной стереоизомерии в дегидрогенизации циклогексановых углеводородов [15], где, кроме секстетного (см. выше), найден дублетный механизм на окисях хрома, молибдена, ванадия и др. (фиг. 4, табл. 4),

В соответствии с опытом следствия из обоих механизмов различны.

При дублетном механизме могут дегидрироваться не только 6-членные циклы [15, 10], катализаторами могут служить не только вещества группы А 1 и А 3, при реакции образуются циклоолефины, дегидрогенизация протекает труднее.

Обмен бензола с дейтерием происходит также на окиси хрома [45], отличаясь от такой же реакции на металлах. Для различения механизмов мог быть также использован [10] коэффициент логарифмической зависимости между константами формулы Аррениуса [4, 26].



Фиг. 5

Дублетная дегидрогенизация углеводородов усиленно разрабатывалась в последнее время [8, 10, 13, 23, 25, 34], потому что она дает вещества, важные для синтеза каучука из нефти. Модель дегидрогенизации бутана в бутилен [25] изображена на фиг. 5.

Было найдено, что над $Cu - Cr_2O_3$ *n*-дивинилбензол образуется с такой же скоростью из *n*-диэтилбензола, как стирол из этилбензола [31] и аналогично при других алкилбензолах [24, 30, 32, 33]. Модель фиг. 6 объясняет эти факты. Кольцо укладывается плоско, боковые цепи находят себе место. Водороды отходят к атомам меди А и А'. Углероды притягиваются к атомам В и В'. Аналогично для более длинных цепей. Для моноалкилбензола модель остается, но с одной боковой цепью [42].

Аналогично может быть рассмотрено углеобразование [1, 10, 77, 38]. Разложение циклогексана передается моделью фиг. 7, где каждый атом С притягивается к атому Ni. Промежуточной формой являются радикалы $CH_2<$ на поверхности, обнаруживаемые по их метилирующему действию [56]. Образующийся на металлах уголь растет в форме дендритов [27, 38].

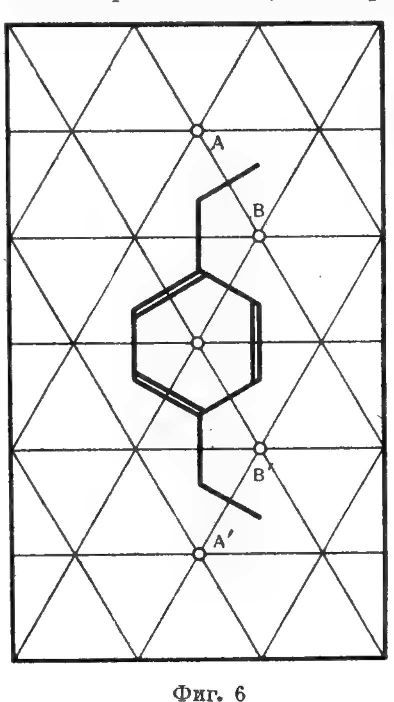
3. Энергетические факторы [2]. Для разрыва связи А—В на катализаторе К требуется меньшая энергия $E_{AB, K}$, чем в газе, Q_{AB}, потому что части А и В связываются дублетом катализатора с выделением энергии

$$E_{\text{AB, K}} = Q_{\text{AB}} - Q_{\text{AK}} - Q_{\text{BK}}. \qquad (1)$$

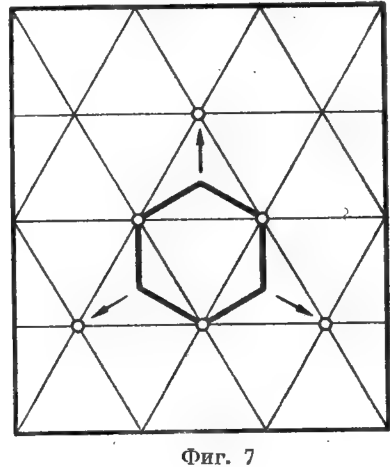
Теплоты образования мультиплетного комплекса (*а* → *б*, фиг. 1), что соответствует активированной адсорбции [79], и его распада (*б* → *в*) суть

$$E' = -E_{\text{OH, Ni}} - E_{\text{CH, Ni}},$$
$$E'' = E_{\text{HH, Ni}} + E_{\text{CO, Ni}}. \qquad (2)$$

Скорость реакции определяется скоростью наиболее эндотермичной элементарной стадии; чем отрицательнее E' (или E''), тем меньше скорость реакции. Во многих случаях входящие в уравнения (1) и (2) значения Q известны из оптических и термохимических данных [77]. Подстановка их дает, что, например, на Ni скорость задается образованием, а на угле — разложением мультиплетного комплекса [37]. Уравнения (1) и (2) позволяют найти расчетом последовательность, в которой располагаются реакции разных типов по их скорости (табл. 2). Расчет производится для разрыва связей, но последовательность скоростей реакции сохраняется и для не зашедшей так далеко деформации связей, т. е. для катализа [7].



Фиг. 6


Фиг. 7

41*

Таблица 2

ПОСЛЕДОВАТЕЛЬНОСТЬ РЕАКЦИЙ НА НИКЕЛЕВОМ КАТАЛИЗАТОРЕ, ВЫЧИСЛЕННАЯ ИЗ $E_{AB,\ K}$

	C—C | | H H	C H | | C H	C—N | | H H	N C | | H N	C—C | | H O	C—C | | H S	C—O | | H H	C—C ‖ | C H
E' =	— 54	— 42	— 42	— 41	— 34	— 32	— 25	— 24
	C H | | N H	C H | | O H	C H | | S H	N H | | Cl H	C H ‖ | C H	C H | | Br H	C H | | Cl H	N H | | O H
E' =	— 23	— 4	— 2	— 1	6	6	9	17

Из табл. 2, в соответствии с опытом, следует, что на никеле дегидрогенизация углеводородов ($E' = -54$) идет труднее, чем их распад с разрывом связей С — С ($E = -42$), который идет наравне с дегидрогенизацией аминов ($E' = -42$) и с реакцией диспропорционирования аминов, например, анилина в дифениламин и аммиак ($E' = -41$). Затем по степени легкости следует дегидратация спиртов ($E' = -34$), далее — отщепление H_2S из меркаптанов ($E' = -32$), затем дегидрогенизация спиртов ($E' = -25$; ср. фиг. 1). Легче всего восстанавливается водородом связь между азотом и кислородом ($E' = 17$) и несколько труднее С — Cl ($E' = 9$ кал/моль). Действительно, оправдалось предсказание теории, что *п*-хлорнитробензол сначала дает *п*-хлоранилин, а затем — анилин [42].

Для гидрогенолиза над никелем, в условиях гидрирования по Сабатье [63], такой же расчет согласуется с опытом более чем для сотни веществ [2] и почти без исключений. Примеры приведены в табл. 3.

Изучение гидридов [22] и формиата [20] никеля показывает, что для получения подобных результатов простая теория промежуточных соединений Сабатье [63] недостаточна, хотя это и не значит, что химические промежуточные соединения никогда не образуются; они образуются, например, при каталитическом разложении формиатов или парофазном гидролизе галоидарилов [64, 65]. Расчеты по мультиплетной теории были выполнены также при изучении реакции бутана с CO_2 над Ni [35] и восстановления галоидопроизводных водородом над углем [37].

4. Классификация. Атомы и связи внутри рамки представляют собой так называемый индекс (см. фиг. 1 или, в общем виде, фиг. 8).

Таблица 3

РЕАКЦИИ РАСЩЕПЛЕНИЯ ПРИ ГИДРОГЕНИЗАЦИИ НАД НИКЕЛЕМ

Класс соединений	Вещество	А—ожидаемые, В—наблюдаемые продукты реакции
Амиды кислот	$CH_3—CO—NH_2$	А. 1) $CH_3·CH_2·NH_2$; H_2O; 2) $CH_3·CH_3$; NH_3; 3) CH_4 В. 1) $CH_3·CH_2·NH_2$; H_2O; 2) $CH_3·CH_3$; NH_3; 3) CH_4
Нитропарафины	$CH_3—CH_2—NO_2$	А. 1) $CH_3·CH_2·NH_2$; H_2O; 2) $CH_3·CH_3$; NH_3; 3) CH_4 В. 1) $CH_3·CH_2·NH_2$; H_2O; 2) $CH_3·CH_3$; NH_3; 3) CH_4
Нитриты	$CH_3—O—NO$	А. 1, 2) CH_3OH; NH_3; H_2O; 3) CH_4; H_2O В. 1, 2) CH_3OH; NH_3; H_2O; 3) CH_4; H_2O
Кетоксимы	$(CH_3)_2C—N—OH$	А. 1) $(CH_3)_2C:NH \rightarrow (CH_3)_2CH·NH_2$; H_2O; 2) $(CH_3)_2CH_2$; NH_3 В. 1) $(CH_3)_2C:NH \rightarrow (CH_3)_2CH·NH_2$; H_2O; 2) $(CH_3)_2CH_2$; NH_3
Амины	$C_6H_5—CH_2—NH_2$	А. 1) $C_6H_5·CH_3$; NH_3 В. 1) $C_6H_5·CH_3$; NH_3
Ароматические спирты	$C_6H_5—CH_2—OH$	А. 1) $C_6H_5·CH_3$; H_2O В. 1) $C_6H_5·CH_3$; H_2O
Арилхлориды	$C_6H_5—Cl$	А. 1) C_6H_6; HCl; 2) CH_4 В. 1) C_6H_6; HCl; 2) CH_4
Эфиры хлорированных кислот	$CH_2Cl—CO—OR$	А. 1) $CH_3·COOR$; HCl В. 1) $CH_3·COOR$; HCl
Бромпроизводные ариламинов	$Br—C_6H_4\ NH_2$	А. 1) $C_6H_5·NH_2$; HBr; 2) C_6H_6; NH_3; 3) CH_4 В. 1) $C_6H_5·NH_2$; HBr; 2) C_6H_6; NH_3; 3) CH_4

Варьируя их, можно получить классификацию реакций в органическом катализе (табл. 4). Так как разным K соответствуют разные Q_{AK} и т. д., то разным индексам должны соответствовать разные катализаторы.

Из табл. 4 видно также, что главные классы реакций органического катализа действительно выражаются при помощи дублетных индексов [6].

Таблица 4

СООТВЕТСТВИЕ МЕЖДУ ГРУППАМИ РЕАКЦИЙ И ГРУППАМИ КАТАЛИЗАТОРОВ

Группа реакций	Индекс	Группа катализаторов
Дегидрогенизация углеводородов	H_2 [C—C / H H] H_2	Cr_2O_3, Mo_2O_3, V_2O_5
Дегидрогенизация спиртов	H_2 [C—O / H H]	Cu, MnO, ZnO, CdO
Дегидратация спиртов	H_2 [C—C / H O] H_2	Al_2O_3, ThO_2, W_2O_5
Разложение спиртов	H_2 [C—O / C H] H_3	C
Гидрогенизация олефинов	H_2 [C H / C H] H_2 (C=C)	Pt, Pd, Ni, Fe, Cu
Кетонизация кислот	H_3 [C C / C O] CH_3, O; HO, O; H	$CaCO_3$, $BaCO_3$, MnO, ThO_2
Хлорирование	H_3 [C Cl / H Cl]	$BiCl_3$, $SbCl_3$, C
Присоединение HCl к олефинам	H_2 [C Cl / C H] H_2 (C=C)	$AlCl_3$, $BaCl_2$
Окисление олефинов	H_2 [C O / C O] H_2 (C=C, O=O) → H_2 [C—O / C—O] H_2	Os, Cu

Следующими по сложности будут реже встречающиеся триплетные реакции и т. д. [6].

Исчерпывающим варьированием была получена полная система дублетных реакций для атомов C, H, N, O, S и Cl, насчитывающая около 2000 типов [5]. Она близка к обычной, но более детализирована: например, при дегидрогенизации углеводородов и спиртов она предусматривает разные катализаторы (см. табл. 4).

Начало классификационной таблицы для реакций гидрогенизации приведено в табл. 5. Первый столбец — это гидрогенолиз (в частности, гидрирование с раскрытием цикла), в остальных случаях первая строка соответствует присоединению водорода к олефинам и ацетиленам, вторая — к иминам и нитрилам, третья—к альдегидам и кетонам.

Таблица 5

ТИПЫ РЕАКЦИЙ КАТАЛИТИЧЕСКОЙ ГИДРОГЕНИЗАЦИИ

$$\begin{array}{ccc}
\begin{array}{cc} C & H \\ | & | \\ C & H \end{array} &
\begin{array}{cc} C & H \\ \| & | \\ C & H \end{array} &
\begin{array}{cc} C & H \\ ||| & | \\ C & H \end{array} \\[2ex]
\begin{array}{cc} C & H \\ | & | \\ N & H \end{array} &
\begin{array}{cc} C & H \\ \| & | \\ N & H \end{array} &
\begin{array}{cc} C & H \\ ||| & | \\ N & H \end{array} \\[2ex]
\begin{array}{cc} C & H \\ | & | \\ O & H \end{array} &
\begin{array}{cc} C & H \\ \| & | \\ O & H \end{array} &
\end{array}$$

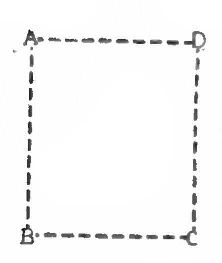
Классификация имеет эвристическое значение; еще не известные реакции, входящие в нее, подлежат открытию (если они не противоречат термодинамике). Так были найдены реакции [7] над Pt и Ni:

$$\begin{array}{l} CH_3{\cdot}CO \\ \qquad H \end{array}\!\!>\!\begin{array}{cc} N & H \\ |\!+ & | \\ Br & H \end{array} = CH_3CONH_2 + HBr; \quad \begin{array}{ccc} (CH_3)_3C & \!\!-\!O & H \\ & |\!+ & | \\ & C & H \end{array} = (CH_3)_3COH + HCl.$$

В индексах I (гидролиз этилового эфира), II (образование амина из спирта и NH_3 или анилина), III (дегидратация спиртов) постепенно заменяется один внутрииндексный атом (левый верхний угол):

$$\mathrm{I}\ \begin{array}{cc} O & C \\ | & | \\ H & O \end{array} \qquad \mathrm{II}\ \begin{array}{cc} N & C \\ | & | \\ H & O \end{array} \qquad \mathrm{III}\ \begin{array}{cc} C\!- & \!C \\ | & | \\ H & O \end{array}$$

Оказалось (6, 47, 67, 68], что прибавление ZnO, SnO, Fe_2O_3—катализаторов дегидрогенизации спиртов — к окиси алюминия ускоряет реакцию I, так как разрывается связь O—H, как при дегидрогенизации спиртов, и замедляет реакции II и III, где нет аналогичного соответствия. Эти тонкие различия могли быть использованы на практике.



Фиг. 8

Важная для синтеза циклических углеводородов реакция Кижнера не укладывалась в мультиплетную теорию. Было экспериментально показано, что на самом деле реакция проходит в две дублетные стадии, принадлежащие к полной системе. Так, циклогексилиденгидразин над Pt переходит в циклогексилдиимин, далее уже распадающийся на циклогексан и азот [16].

В известных до сих пор случаях дегидрогенизации аминов выделялось несколько молекул H_2 (сложная реакция). Простая дублетная дегидрогенизация, предвидимая мультиплетной теорией, найдена недавно автором и Н. А. Васюниной на примере превращения 1-диэтиламино-4-аминопентана в 1-диэтиламино-4-иминопентан над Pd, Pt и Cu [19].

Индексы мультиплетной теории играют роль функциональных групп в классической структурной теории [5, 6].

Во вторую очередь необходимо учитывать влияние заместителей, например, при A (фиг. 8), на скорость.

Из уравнений мультиплетной теории следует [12], что

$$\Delta E' = \Delta Q_{AK} - \Delta Q_{AB}, \qquad (3)$$
$$\Delta E'' = \Delta Q_{AD} - \Delta Q_{AK}, \qquad (4)$$

т. е. если замещение повышает скорость реакции, то это обязано тому, что прирост энергии связи с катализатором превышает таковой для реагирующей связи [уравнение (3), ср. § 6].

5. К и н е т и к а. Мультиплетная теория тесно связана с кинетикой. Дублетный механизм (фиг. 9), дополненный представлением о том, что H_2 диссоциирует на атомы H на одних парах атомов металла (K = 2A) и гидрирует вещество B, адсорбированное на других парах атомов K', причем

процесс проходит через полугидрированную форму X, приводит к теории гидрогенизации [11], количественно описывающей обширный опытный материал. Схема реакции имеет вид (C — гидрированное вещество):

$$\text{I. } B + K' \overset{b'}{\rightleftharpoons} B.K'; \quad H_2 + K' \overset{h'}{\rightleftharpoons} H_2 \cdot K'; \quad C + K' \overset{c'}{\rightleftharpoons} C.K',$$

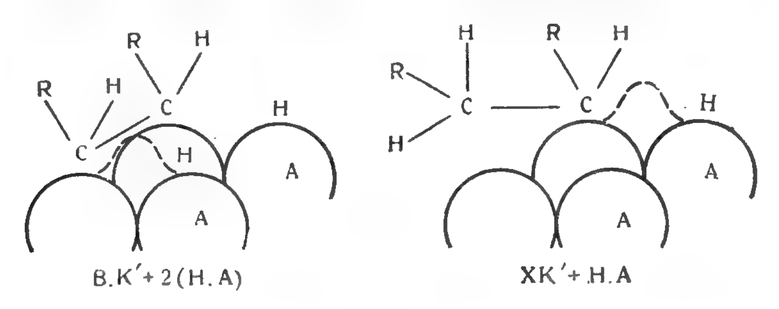
$$\text{II. } B + K \overset{b}{\rightleftharpoons} B.K; \quad H_2 + K \overset{h}{\rightleftharpoons} H_2K; \quad H_2.K \overset{z}{\rightleftharpoons} 2\,(H.A); \quad C + K \overset{c}{\rightleftharpoons} C.K,$$

$$\text{III. } BK' + H.A \underset{i}{\overset{q}{\rightleftharpoons}} X.K'.A; \quad X.K'.A + H.A \underset{j}{\overset{f}{\rightleftharpoons}} C.K' + K.$$

I и II — адсорбционные равновесия, b' и т. д. — их константы. Из схемы можно вывести, что скорость гидрогенизации

$$v_h = k_h \frac{b'[B]}{1 + b'[B] + c'[C] + h'[H_2]} \; \frac{h[H_2]}{1 + b[B] + c\,[C] + h\,[H_2]} \tag{5}$$

(где $k_h = qfzss'/i$, s и s' — числа активных центров) и аналогично для дегидрогенизации и обмена с дейтерием. Теория показывает, что уравнение

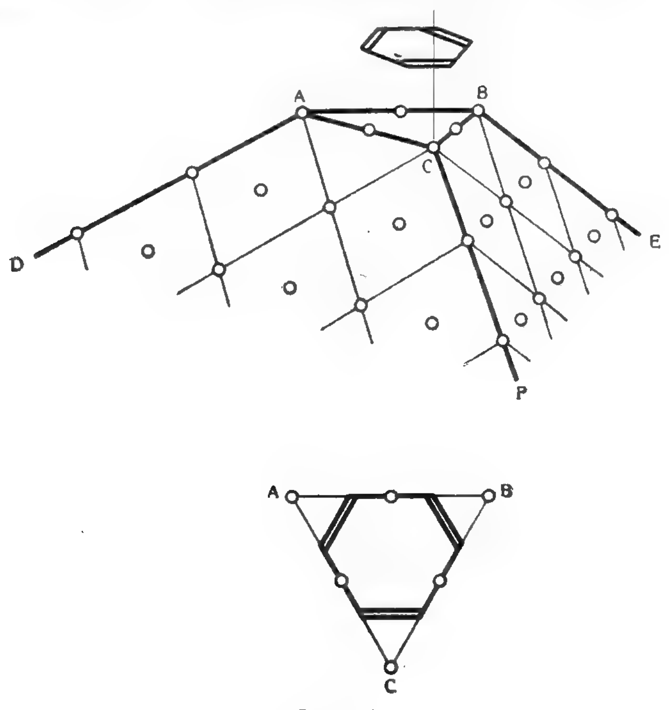


Фиг. 9

(5) верно и для бензола. На модели (рис. 10) центры K принадлежат ребрам, где Ni — Ni = 3.52Å, а центр K' — грани (111), где Ni — Ni = 2.49Å, в согласии с тем, что по Эйрингу [73] при 3.52Å энергия активации диссоциативной адсорбции H_2 на дублете Ni меньше, чем при 2.49Å. Опытами автора [40, 46, 49, 54] найдено, что $b' \approx c'$; это говорит об одинаковой ориентации B и C на центрах K' [50]. В предельных случаях отдельные члены в уравнении (5) выпадают и уравнение упрощается.

Теория объясняет нулевой порядок и по B и по H_2 при гидрогенизации, наблюдающийся при больших концентрациях, а также нулевой порядок по B в начале и первый — в конце реакции. Изменение адсорбционных

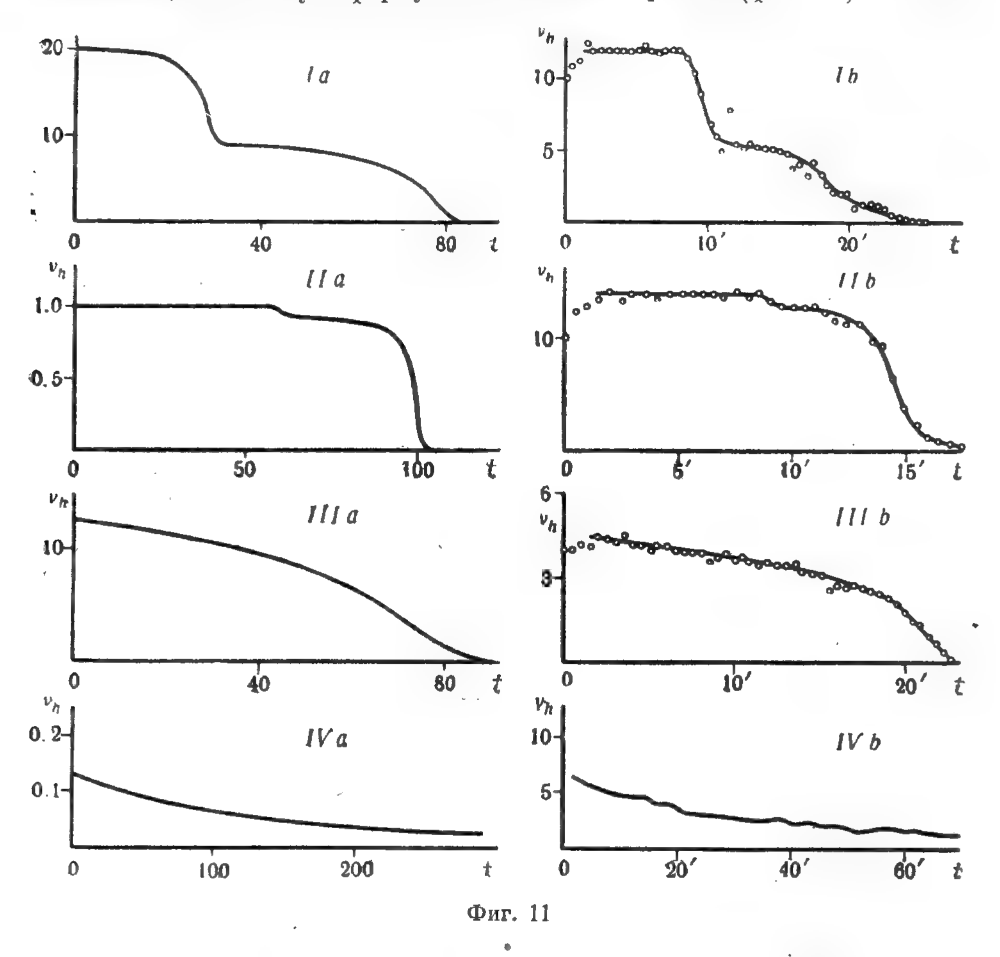
коэффициентов и констант скоростей с температурой в уравнении (5) по обычному экспоненциальному закону объясняет наличие первого порядка по H_2 при высоких и нулевого порядка при низких температурах, обращение температурного коэффициента скорости и влияние давления на последний. Разность теплот адсорбции H_2 и B (толуола) на центрах



Фиг. 10

K' находится равной 26 500 кал/моль. Уравнение (5) правильно передает уменьшение v_h и повышение порядка реакции при переходе от гидрогенизации в паровой к гидрогенизации в жидкой фазе вследствие появления в знаменателях членов, относящихся к растворителю. Аналогично объясняется влияние примеси CO. Теория разработана для ряда катализаторов (Ni, Pt, Cu). Она согласуется с замедлением пара-орто-превращения H_2 олефинами, передает правила Лебедева-Вавона для гидрогенизации, применима к последовательному гидрированию (например, производных ацетилена).

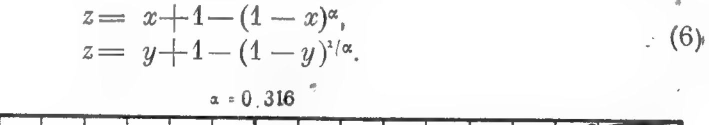
Особенное значение новая теория приобретает для гидрогенизации смесей; она впервые дает [из уравнения (5)] уравнения для этих случаев, описывающих сложную форму кинетических кривых (фиг. 11): слева —



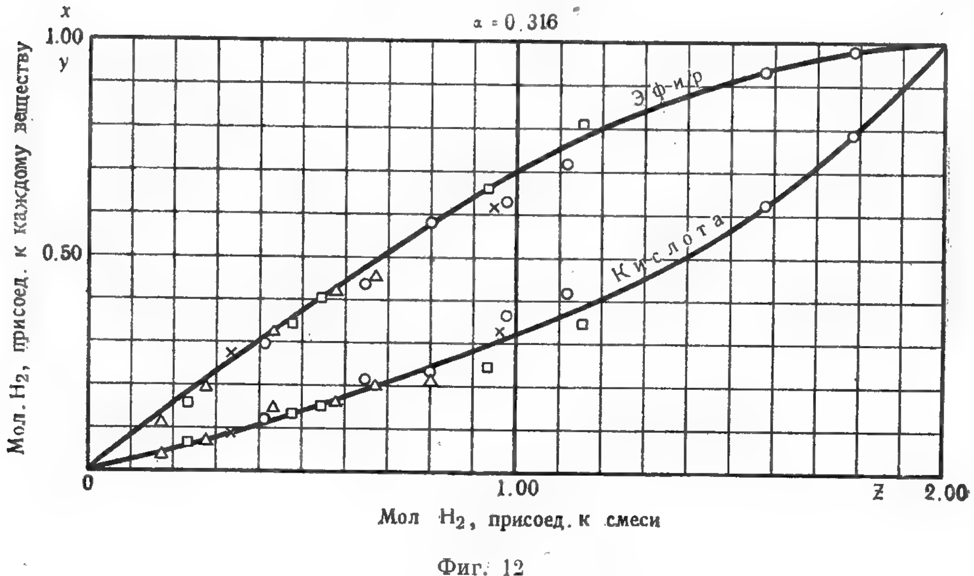
Фиг. 11

вычисленные с произвольными константами, справа — наблюденные по Лебедеву [57] и др.; Ib — смесь эйгенол + метил-трет. бутилэтилен, IIb — метил-трет. бутилэтилен + окись мезитила, IIIb — итаконовая кислота + несимм. дифенилэтилен, IVb — метилдиэтилэтилен + метилдифенилэтилен. При этом различается ступенчатость и избирательность и

указываются условия появления последних. Вводится новая величина — показатель избирательности, $\alpha = k_2b'_2/k_1b'_1$, где значки 1 и 2 относятся к 1 и 2 веществу. При помощи α можно рассчитывать глубины гидрирования каждого компонента x и y как функцию суммарной глубины z



$$z = x + 1 - (1 - x)^{\alpha},$$
$$z = y + 1 - (1 - y)^{1/\alpha}. \qquad (6)$$



Фиг. 12

На фиг. 12 приведен пример для коричных кислоты и эфира; кривые — вычисленные по уравнению (6) при $\alpha = 0{,}316$, точки — наблюденные. Для α выводятся новые закономерности; так, в определенных условиях α связано с изменением свободной энергии

$$RT \ln \alpha \approx \Delta F_1^0 - \Delta F_2^0. \qquad (7)$$

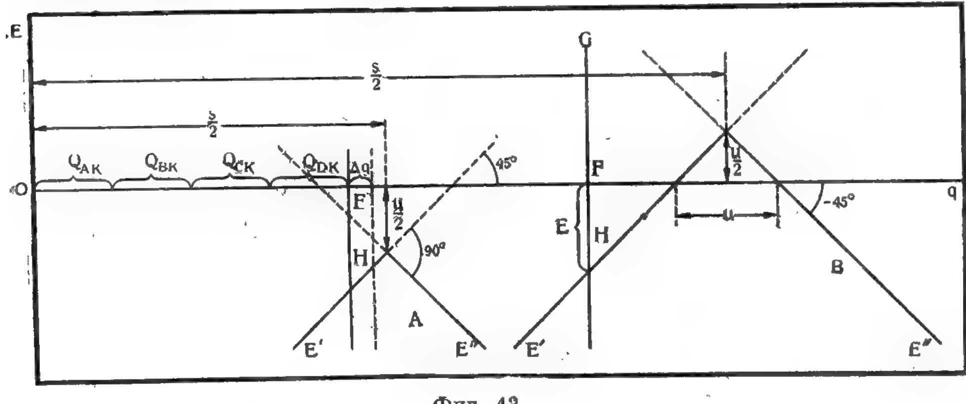
Другие работы автора по каталитической кинетике — см. [3, 14, 28, 36, 52].

6. Активность и подбор катализатора [12]. Этот вопрос — центральный в катализе. Для реакции AB + CD = AD + BC на катализаторе K (фиг. 8) обозначим теплоту реакции через $u = -Q_{AB} - Q_{CD} + Q_{DA} + Q_{BC}$, сумму энергий разрываемых и образующихся связей

через $s = Q_{AB} + Q_{CD} + Q_{AD} + Q_{BC}$ и адсорбционный потенциал через $q = Q_{AK} + Q_{BK} + Q_{CK} + Q_{DK}$. Тогда уравнения типа уравнения (1) дают

$$\text{I.} \quad E' = q - \frac{s}{2} + \frac{u}{2}, \qquad \text{II.} \quad E' = -q + \frac{s}{2} + \frac{u}{2}, \tag{8}$$

где I относится к адсорбции, II — к десорбции. Реакцию задерживает тот энергетический барьер E' или E'', который больше E. Деформационный характер катализа [53] учитывается условием $\varepsilon = \gamma E$, где ε — энергия активации, $\gamma \approx 3/4$. Уравнения (8) выражают функциональную зависимость между адсорбцией и катализом и графически изображены на фиг. 13 (А для эндо- и В для экзотермической реакции); FH — энергетические барьеры. Как видно, оптимальные условия подбора катализатора (при А) состоят в равенстве адсорбционного потенциала q среднему из значений энергии разрываемых и образующихся связей $s/2$; при этом энергетический барьер E равен половине теплового эффекта каталитической реакции $u/2$. Итак, в катализе, кроме структурного, должно существовать энергетическое соответствие. Слишком слабая и слишком сильная адсорбция вредит катализу.



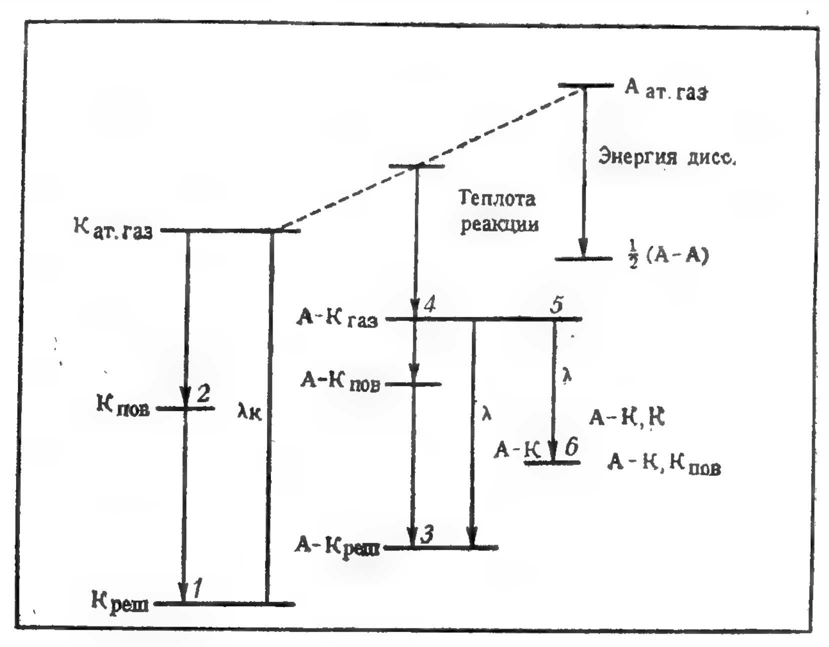
Фиг. 13

Можно повысить активность катализатора с данным $q = q_1$, изменяя слагаемые Q_{AK} и т. д. так, чтобы q_1 ближе подходило к условию оптимума

$q = s/2$. Поэтому, если $q_1 < s/2$, то скорость реакции возрастает (фиг. 13), когда

$$0 < \Delta q < s - 2q_1. \tag{9}$$

Первым способом для этого является замена одного из атомов К (фиг. 8) другим, $К_1$; сюда относится применение смешанных катализаторов, где



Фиг. 14

реакция проходит на границе раздела твердых фаз. Условие (9) есть количественная формулировка вывода, сделанного из мультиплетной теории еще Швабом [66]. Разница в $Q_{АК}$ и $Q_{АК_1}$ может быть значительной, так $Q_{О-Cu} = 47$, а $Q_{О-Fe} = 62$ ккал/моль.

Другой способ повышения активности состоит в изменении числа атомов, соседних с активным центром на поверхности, — изменение микрошероховатости, или ненасыщенности, о которой можно отчасти судить также на основании термограмм [39, 61]. Мы имеем

$$Q_{А-К}\,(\text{пов.}) = Q_{А-К}\,(\text{реш.}) + \lambda_1, \tag{10}$$

и аналогично для Q_{B-K} и т. д. Из фиг. 14 видно, что

$$\lambda_1 = \alpha\lambda_K - \lambda_{A-K} + \lambda_{A-K,\,K}. \quad (11)$$

Чтобы перейти от Q_{A-K} (реш.) к Q_{A-K} (пов.), атомы K должны пройти путь 1—2, а связи A — K путь 3—4—5—6. A — K, K есть связь A—K в мультиплетном комплексе, гда **K** связано с решеткой; λ_k — энергия сублимации **K** (так, $\lambda_{Ni} = 101$, $\lambda_{NiO} = 117$ ккал/моль). Для гладкой поверхности, по теореме Лапласа, $\alpha = \frac{1}{2}$, для выступов $\frac{1}{2} < \alpha < 1$. Места с большим $\alpha\lambda_k$ являются местами с повышенной энергией и адсорбционным потенциалом.

Генезис влияет на активность катализатора, влияя на число соседей, т. е. на α и λ_1, и изменяя по уравнению (10) Q_{A-K} и т. д., а отсюда по уравнению (8) и E. При $q_1 < \frac{s}{2}$ для возрастания активности надо, чтобы λ_1 возрастало (как в теории Тэйлора [78]), но оставалось в пределах

$$0 < \Delta\lambda < s - 2q_1. \quad (12)$$

Повидимому, аналогично действует носитель [18].

Опыт подтверждает, что активные центры находятся на границе раздела твердых фаз в смешанных катализаторах $Ni + Al_2O_3$ и $ZnO + Al_2O_3$ [41, 17]. При дегидрогенизации несимм. дифенилэтана активные центры находятся на выступах, так как иначе в дублетной модели вторая фенильная группа стерически препятствовала бы реакции [29]. Там же должны находиться активные центры для других алкилбензолов, так как выходы мало отличаются [31]. Эти случаи, а также рассмотренные выше модели (фиг. 2, 6, 7, 10) являются применением мультиплетной теории к исследованию местонахождения и структуры активных центров наложением молекул сложной структуры. Мультиплетная теория считает, что каталитически активный центр — метастабильное образование, не успевший разрастись зародыш кристаллизации [1]. Каталитически активных центров на поверхности меньше, и они более однородны, чем адсорбционные центры [3].

ЛИТЕРАТУРА

1. *А. А. Баландин.* ЖРХО 61, 909, 1929.
2. *А. А. Баландин.* ЖРХО 62, 793, 1930; ЖОХ 12, 337, 1942.
3. *А. А. Баландин.* Труды III ф.-х. конф., Л., 1930, стр. 193; Z. phys. Ch. B. 9, 49, 319, 1930; ЖФХ 4, 268, 1933; ЖОХ 12, 150, 160, 1942.

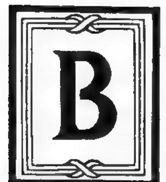
4. *А. А. Баландин.* ЖФХ 4, 257, 1933.
5. *А. А. Баландин.* ЖФХ 5, 679, 1934.
6. *А. А. Баландин.* Усп. химии, 4, 1004, 1935; Успехи катализа в теории и практике, М., 1936.
7. *А. А. Баландин.* ЖФХ 14, 1160, 1940.
8. *А. А. Баландин.* ЖФХ 14, 1378, 1940.
9. *А. А. Баландин.* Успехи химии, 10, 262, 1941; Вестн. АН СССР, № 5—6, 79, 1946.
10. *А. А. Баландин.* ИОХН I, 21, 1942.
11. *А. А. Баландин.* ИОХН 6, 477, 1943; I, 89, 1944; 4, 339, 1945; ЖОХ 15, 608, 619, 770, 781, 1945; 16, 557, 1946.
12. *А. А. Баландин.* ЖОХ 16, 793, 1946.
13. *А. А. Баландин, О. К. Богданова, А. П. Щеглова, М. Н. Марушкин.* ЖПХ 18, 609, 1945.
14. *А. А. Баландин, А. Х. Борк.* Уч. зап. МГУ, 2, 217, 1934.
15. *А. А. Баландин, И. И. Брусов.* ЖОХ 7, 18, 1937.
16. *А. А. Баландин, Д. Н. Васкевич.* ЖОХ 6, 1870, 1878, 1936.
17. *А. А. Баландин, В. Э. Вассерберг.* ИОХН 5, 456, 1945.
18. *А. А. Баландин, Н. А. Васюнина.* ДАН СССР 52, 139, 1946.
19. *А. А. Баландин, Н. А. Васюнина.* ДАН СССР 56, 265, 1947.
20. *А. А. Баландин, Е. С. Григорян, З. Q. Янышева.* ЖОХ 10, 1031, 1940.
21. *А. А. Баландин, Н. П. Егорова.* ДАН СССР 56, 1947.
22. *А. А. Баландин, В. В. Ерофеев, К. А. Печерская, М. С. Стаханова.* ЖОХ 11, 577, 1941; ЖОХ 12, 168, 1942.
23. *А. А. Баландин, Н. Д. Зелинский. О. К, Богданова, А. П. Щеглова.* ЖПХ 14, 435, 1941; 15, 123, 1942; Каучук и резина, № 4, 4, 1941; ИОХН 5, 497, 1946.
24. *А. А. Баландин, Г. М. Марукян, О. К. Богданова.* ЖПХ 14, 161, 1941.
25. *А. А. Баландин, М. Н. Марушкин, Л. Н. Павлов.* ДАН 32, 136, 1941.
26. *А. А. Баландин, Ф. Л. Козьмин.* Уч. зап. МГУ 71, 210, 1941.
27. *А. А. Баландин, Н. З. Котелков.* ЖПХ 15, 139, 1942.
28. *А. А. Баландин, А. Л. Либерман.* ДАН СССР 28, 796, 1940.
29. *А. А. Баландин, Г. Р. Лурье, Г. М. Марукян.* ДАН СССР 53, 131, 1946.
30. *А. А. Баландин, Г. М. Марукян.* ЖПХ 19, 207, 623, 1277, 1946.
31. *А. А. Баландин, Г. М. Марукян.* ДАН СССР, 55, 127, 219, 1947.
32. *А. А. Баландин, Г. М. Марукян, Р. Г. Сеймович.* ДАН СССР 41, 71, 1943.
33. *А. А. Баландин, Г. М. Марукян, А. А. Толстопятова.* ЖПХ 19, 1079, 1946.
34. *А. А. Баландин, М. Н. Марушкин.* ДАН СССР 40, 254, 1943.
35. *А. А. Баландин, М. Н. Марушкин, М. М. Афанасьев.* ИОХН I, 79, 1941.
36. *А. А. Баландин, М. Н. Марушкин, В. А. Иконников.* Уч. зап. МГУ, 2, 221, 1934.
37. *А. А. Баландин, В. В. Патрикеев.* ЖОХ 11, 22, 1941.
38. *А. А. Баландин, В. В. Патрикеев.* ДАН 34, 96, 1942; 40, 171, 1943; ЖОХ 14, 57, 1944.
39. *А. А. Баландин, Т. В. Роде.* Сб. Конф. метод. иссл. кат., М., 1947.
40. *А. А. Баландин, А. М. Рубинштейн.* ЖФХ 5, 12, 1934.
41. *А. А. Баландин, А. М. Рубинштейн.* ЖФХ 6, 576, 1935; JACS 57, 1147, 1935.
42. *А. А. Баландин, А. Н. Титова.* Уч. зап. МГУ, 2, 229, 1934.

43. *А. А. Баландин, А. А. Толстопятова.* ЖОХ 17, печ., 1947.
44. *А. А. Баландин, В. С. Федоров.* ДАН СССР 30, 21, 1941.
45. *А. А. Баландин, З. В. Харина, Р. В. Тейс.* ДАН СССР 53, 719, 1946.
46. *А. А. Баландин, Н. И. Шуйкин.* ЖФХ 5, 707, 1934.
47. *А. А. Баландин, М. П. Несвижский, И. К. Козминская.* ЖОХ 2, 60, 1932.
48. *А. А. Баландин, Я. Т. Эйдус.* ДАН СССР 49, 655, 1945; Усп. химии, 15, 15, 1946.
49. *А. А. Баландин, Ю. К. Юрьев.* ЖФХ 5, 393, 1934.
50. *А. Х. Борк.* ЖФХ 8, 709, 1936.
51. *А. Х. Борк.* ЖФХ 14, 1288, 1940.
52. *А. Х. Борк, А. А. Баландин.* ИОХН 5, 465, 1945.
53. *Н. Д. Зелинский.* Избр. труды, т. 2, М., 1941.
54. *Н. Д. Зелинский, А. А. Баландин.* ИАН СССР 29, 1929.
55. *Н. Д. Зелинский, А. А. Баландин, А. М. Рубинштейн.* Усп. химии, 11, 99, 1942.
56. *Н. Д. Зелинский, Н. И. Шуйкин.* ДАН СССР 3, 255.
57. *С. В. Лебедев.* Жизнь и труды, Л., 1938.
58. *И. Лэнгмюр.* Нов. идеи в химии, № 8, Л., 1924, стр. 38.
59. *Д. И. Менделеев.* ЖРХО 18, 8, 1886; *А. А. Баландин.* Усп. химии, 13, 365, 1944
60. *М. С. Платонов.* Тезисы Катал. конфер. Изд. АН СССР, М., 1940, стр. 35.
61. *Т. В. Роде, А. А. Баландин.* ИОХН 2, 209, 1946.
62. *А. М. Рубинштейн.* ИОХН 5, 509, 1945.
63. *П. Сабатье.* Катализ в органической химии, Л., 1932.
64. *Л. Х. Фрейдлин, А. А. Баландин, А. И. Лебедева.* ИОХН 2, 268, 1941.
65. *Л. Х. Фрейдлин, А. А. Баландин, Г. А. Фридман, А. И. Лебедева.* Химич. пром. № 12, 8, 1945.
66. *Г. М. Шваб.* Катал. с точки зрения химич. кинетики, Л., 1934, стр. 185.
67. *Н. И. Шуйкин, А. А. Баландин, Ф. Т. Дымов.* ЖОХ 4, 1452, 1934.
68. *Н. И. Шуйкин, А. А. Баландин, З. И. Плоткин.* ЖОХ 4, 1451, 1934.
69. *O. Beeck.* Rev. Mod. Phys. 17, 79, 1945; *A. Wheeler, A. E. Smith.* Proc. Roy. Soc. 2, 177, 62, 1940.
70. *S. Berkman, J. C. Morrell, G. Egloff.* Catalysis inorg. and organic, N. Y., 1940, p. 41.
71. VIII Catalysis Report. J. Phys. Chem., 34, 2129, 1930.
72. *F. H. Constable.* Proc. Roy. Soc. A, 107, 279, 1925; 108, 355, 1925; 110, 283, 1926.
73. *S. Glasstone. K. I. Laidler, H. Eyring.* Theory of Rate Processes, N. Y. a. London 1943, p. 345; готовится русск. пер. под ред. А. А. Баландина.
74. *R. K. Greenhalgh, M. Polanyi.* Tr. Far. Soc. 35, 511, 1939; *G. H. Twigg, E. K. Rideal.* Proc. Roy. Soc. A, 171, 55, 1939.
75. *P. H. Emmett, N. Skau.* JACS 65, 1929, 1943.
76. *J. H. Long, I. C. W. Frazer, E. Ott.* JACS 56, 1101, 1934.
77. *L. Pauling.* Nature of the Chemical Bond, N. Y., 1944, p. 409.
78. *H. S. Taylor.* J. Phys. Chem. 30, 145, 1926; XII Catalysis Report, N. Y., 1941, p. 41.
79. *H. S. Taylor.* JACS 53, 578, 1931; *M. Polanyi.* J. Soc. Chem. Ind. 54, 123, 1935.
80. *H. S. Taylor.* Amer. Scientist 34, 553, 1946.

Академик
А. Н. Несмеянов,
Р. Х. Фрейдлина, А. Е. Борисов

ХИМИЯ КВАЗИКОМПЛЕКСНЫХ МЕТАЛЛООРГАНИЧЕСКИХ СОЕДИНЕНИЙ И ЯВЛЕНИЯ ТАУТОМЕРИИ

ВВЕДЕНИЕ



В течение последних лет мы (в некоторой части работ совместно с И. Ф. Луценко, Н. С. Кочетковой, Э. М. Брайниной, А. К. Кочетковым, О. В. Ногиной, Э. И. Кан) изучали химию таких металлоорганических соединений, которые могут рассматриваться как продукты, прямо и косвенно образованные присоединением солей металлов (Hg, Sb, Sn, Pb, Tl, а также J) к непредельным молекулам, подобным ацетилену и этилену и их производным и окиси углерода.

Вещества этого класса замечательны тем, что в своих химических реакциях они ярко проявляют двойственность, похожую на двойственность реагирования таутомерных соединений. Они то реагируют как истинные металлоорганические вещества, образованные путем присоединения по кратной связи непредельной молекулы, то вступают в реакции как комплексные аддукты металлических солей и ацетилена (соотв. олефинов, окиси углерода), в такие реакции, в которых эти непредельные адденды подвергаются замещению. В соответствии с этим такие вещества обозначены нами как квазикомплексные. При ближайшем изучении этих соединений, однако, становится совершенно ясным, что речь идет о чем-то существенно ином, чем классическое явление таутомерии, предсказанное Бутлеровым [1] на основе развитой им теории строения и сводящееся к

равновесной обратимой реакции взаимного превращения структурных изомеров. Эта двойственность реакций веществ данного типа наиболее ярко может быть проиллюстрирована на подробно изученных нами четырех производных ацетилена и сулемы, два из которых были известны и ранее [2, 3], но оставались едва изученными, а два открыты нами.

КВАЗИКОМПЛЕКСНЫЕ АДДУКТЫ АЦЕТИЛЕНА

Если пропускать ацетилен в 15% водную соляную кислоту, насыщенную сулемой [4], то выпадает вещество состава $C_2H_2HgCl_2$ с т. пл. 124° и дипольным моментом 1.64 D, для которого описанные ниже реакции, несомненно, доказывают структуру β-хлорвинилмеркурхлорида ClCH=CHHgCl. Это вещество при обработке аммиаком, лучше всего в растворе бензола, превращается в ди-β-хлорвинилртуть с т. пл. 70°, в свою очередь при действии сулемы переходящую в исходный β-хлорвинилмеркурхлорид с т. пл. 124°.

При присоединении паров сулемы к ацетилену при 120° [5], а также при описанной ниже реакции обмена три-цис-β-хлорвинилсурьмы с сулемой образуется стереоизомер β-хлорвинилмеркурхлорида с т. пл. 79° и дипольным моментом 2.76 D [6, 7], который, в свою очередь, при обработке аммиаком переходит в жидкий стереоизомер ди-β-хлорвинилртути (т. к. при $1.3 \cdot 10^{-3}$ мм 52.5—53°, $d\,^{20°}_{4°} = 2.7929$, $n\,^{20°}_{D} = 1.6124$, $MR = 40.23$, $\sigma^{21°} = 47.18$ эрг/см², $P_{найд} = 304.4$, $P_{выч} = 309.3$), способный перегоняться в глубоком вакууме.

Это вещество при действии сулемы образует опять исходный β-хлорвинилмеркурхлорид с т. пл. 79° [8]. Оба ряда стереоизомеров не превращаются друг в друга ни при хранении в темноте чистых веществ, ни при хранении их растворов, но твердая ди-β-хлорвинилртуть легко превращается в свой жидкий стереоизомер при облучении светом ртутной лампы [8], что обычно для превращения трансизомеров в цисизомеры. Эта жидкая ди-β-хлорвинилртуть — чистый цис-цисизомер, так как при действии сулемы и при действии хлора

$$(Cl\,CH{=}CH)_2\,Hg + Hg\,Cl_2 \rightarrow 2Cl\,CH{=}CH\,Hg\,Cl, \qquad (1)$$

$$(Cl\,CH{=}CH)_2\,Hg + Cl_2 \rightarrow Cl\,CH{=}CH\,Hg\,Cl + Cl\,CH{=}CH\,Cl \qquad (2)$$

она дает исключительно соответствующий ей стереоизомер β-хлорвинилмеркурхлорида с т. пл. 79°, подобно тому как ди-β-хлорвинилртуть

42*

с т. пл. 70° дает при подобных же реакциях тоже исключительно β-хлорвинилмеркурхлорид с т. пл. 124° [8]. Нет сомнений, что низкоплавкий (79°) стереоизомер с большим дипольным моментом (2.76 D) — цисизомер, а вещество с т. пл. 124° — транссоединение, как это следует также из сопоставления по правилу Вернера точек плавления обоих изомеров с точками плавления более низкоплавкого ортохлорфенилмеркурхлорида и его высокоплавкого параизомера [7]. Это следует, кроме того, и из рентгенографического исследования транссоединения, произведенного по нашему предложению А. И. Китайгородским [9].

Все эти четыре вещества несомненно имеют строение β-хлорвинильных ртутноорганических веществ, так как в целой серии обменных реакций они дают хлорвинильные продукты превращения:

$$\mathrm{Cl\,CH{=}CH\,Hg\,Cl + J_2 \rightarrow Cl\,CH{=}CHJ + Hg\,ClJ} \quad (3)$$

$$\mathrm{Cl\,CH{=}CH\,Hg\,Cl + Cl_2 \rightarrow Cl\,CH{=}CH\,Cl + HgCl_2} \quad (4)$$

$$\mathrm{Cl\,CH{=}CH\,Hg\,Cl + HBr \rightarrow Cl\,CH{=}CH_2 + Hg\,Cl\,Br} \quad (5)$$

$$\mathrm{(Cl\,CH{=}CH)_2\,Hg + JCl_3 \rightarrow Cl\,CH{=}CHJ\,Cl_2 + Cl\,CH{=}CH\,Hg\,Cl} \quad (6)$$

$$\mathrm{(Cl\,CH{=}CH)_2\,Hg + ClCH{=}CHJCl_2 \rightarrow (Cl\,CH{=}CH)_2\,JCl + Cl\,CH{=}CH\,Hg\,Cl} \quad (7)$$

$$\mathrm{(Cl\,CH{=}CH)_2\,Hg + Sn\,Cl_2 \rightarrow (Cl\,CH{=}CH)_2\,Sn\,Cl_2 + Hg} \quad (8)$$

$$\mathrm{(Cl\,CH{=}CH)_2\,Hg + Pb\,(OOCCH_3)_4 \rightarrow (Cl\,CH{=}CH)_2\,Pb\,(OOCCH_3)_2 + Hg\,(OOCCH_3)_2.} \quad (9)$$

В ряде других реакций эти соединения с чрезвычайной легкостью элиминируют ацетилен, отщепляя элементы сулемы. Элиминация ацетилена происходит для дихлорвинильных соединений по ступеням при нагревании выше точки плавления наиболее высокоплавкого из четырех веществ (124°):

$$\mathrm{(Cl\,CH{=}CH)_2\,Hg \rightarrow Cl\,CH{=}CH\,Hg\,Cl + C_2\,H_2,} \quad (10)$$

$$\mathrm{Cl\,CH{=}CH\,Hg\,Cl \rightarrow Hg\,Cl_2 + C_2H_2.} \quad (11)$$

Еще легче, уже при комнатной температуре, элиминация ацетилена происходит при действии на все четыре вещества многих реагентов, способных тем или иным способом связать сулему, например, в форме комплексного соединения. Так реагирует водный раствор иодистого калия, цианистого калия, тиосульфата натрия и т. п. Эти реакции чрезвычайно напоминают «реакции внедрения» Вернера, т. е. реакции обмена

аддендов первой сферы комплекса, столь типичные для химии комплексных соединений, и их с этой точки зрения можно изобразить следующим образом:

$$C_2H_2Hg\genfrac{}{}{0pt}{}{Cl}{Cl}+3KJ \rightarrow C_2H_2+[JHg\genfrac{}{}{0pt}{}{J}{J}]K+2KCl, \qquad (12)$$

$$\genfrac{}{}{0pt}{}{C_2H_2}{C_2H_2}Hg\genfrac{}{}{0pt}{}{Cl}{Cl}+4KJ \rightarrow 2C_2H_2+[\genfrac{}{}{0pt}{}{J}{J}Hg\genfrac{}{}{0pt}{}{J}{J}]\genfrac{}{}{0pt}{}{K}{K}+2KCl, \qquad (13)$$

$$\genfrac{}{}{0pt}{}{C_2H_2}{C_2H_2}Hg\genfrac{}{}{0pt}{}{Cl}{Cl}+4KCN \rightarrow 2C_2H_2+[\genfrac{}{}{0pt}{}{CN}{CN}Hg\genfrac{}{}{0pt}{}{CN}{CN}]\genfrac{}{}{0pt}{}{K}{K}+2KCl, \qquad (14)$$

$$C_2H_2HgCl_2+2Na_2S_2O_3 \rightarrow C_2H_2+Na_2[Hg(S_2O_3)_2]+2NaCl. \qquad (15)$$

Очень наглядна в этом смысле реакция вытеснения ацетилена трифенилфосфином

$$C_2H_2HgCl_2+2P(C_6H_5)_3 \rightarrow C_2H_2+\genfrac{}{}{0pt}{}{Cl}{Cl}Hg\genfrac{}{}{0pt}{}{P(C_6H_5)_3}{P(C_6H_5)_3} \qquad (16)$$

$$2C_2H_2HgCl_2+2P(C_6H_5)_3 \rightarrow 2C_2H_2+\genfrac{}{}{0pt}{}{Cl}{Cl}Hg\genfrac{}{}{0pt}{}{P(C_6H_5)_3}{P(C_6H_5)_3} \qquad (17)$$

К таким же реакциям, приводящим к полной элиминации ацетилена, относятся реакции алкилирования (или, точнее, арилирования) посредством оловоорганических соединений

$$C_2H_2HgCl_2+2C_6H_5SnCl_3 \xrightarrow{NaOH} (C_6H_5)_2Hg+C_2H_2+Na_2SnO_3, \qquad (18)$$

$$(C_2H_2)_2HgCl_2+2C_6H_5SnCl_3 \xrightarrow{NaOH} (C_6H_5)_2Hg+2C_2H_2+Na_2SnO_3, \qquad (19)$$

гриньярова реактива и реакции хлорметилирования

$$C_2H_2HgCl_2+CH_2N_2 \rightarrow ClCH_2HgCl+N_2+C_2H_2. \qquad (20)$$

Во всех этих реакциях оба аддукта ацетилена ведут себя как свободная сулема.

При действии окиси серебра в воде на транс-β-хлорвинилмеркурхлорид он может быть превращен в растворимую гидроокись $ClCH=CHHgOH$, а последняя в ряд солей [10] $ClCH=CHHgBr$ (т. пл. 121—122°), $ClCH=CHHgCN$ (т. пл. 101—102°), $ClCH=CHHgOOCCH_3$ (т. разл. 85°),

$ClCH=CHHgOOCC_6H_5$ (т. пл. 120—122°), $ClCH=CHHgC\equiv CC_6H_5$ (т. пл. 102—103°), фталимидное производное — $ClCH=CHHgN(CO)_2C_6H_4$ (т. пл. с разл. 166—167°). Соли, в частности иодид, могут также быть получены гладко идущей реакцией дихлорвинилртути с соответствующей солью ртути [11] $(ClCH=CH)_2Hg+HgJ_2\rightarrow 2ClCH=CHHgJ$ (трансизомер, т. пл. 114—115.5°). Последние реакции, так же как широкий ряд обменов от (1) до (9), и также отношения цис-трансизомерии не оставляют сомнения в строении описываемых веществ как хлорвинильных ртутноорганических соединений, тогда как ряд реакций (10) — (20) мог бы соответствовать реакциям молекулярных аддуктов сулемы и ацетилена. Этот характер молекулярного аддукта выявляется и при рентгенографическом исследовании транс-хлорвинилмеркурбромида (А. И. Китайгородский). Это вещество кристаллизуется в ромбической сингонии (группа v_h^{16}) с размерами ячейки $a=6.84$, $b=18.42$, $c=4.17$. HgClBr кристаллизуется аналогично, причем размеры ячейки $a=6.78$, $b=13.17$, $c=4.10$.

Элементарная ячейка хлоробромида ртути как бы раздвинулась по одному направлению, впустив молекулу ацетилена, причем два остальных параметра и весь характер симметрии кристалла остались неизменными. Перед нами, однако, не новый вид таутомерии

$$ClCH=CHHgCl \rightleftarrows C_2H_2\cdot Hg\begin{matrix}Cl\\ Cl\end{matrix} \qquad (21)$$

$$(ClCH=CH)_2Hg \rightleftarrows \begin{matrix}C_2H_2\\ C_2H_2\end{matrix}Hg\begin{matrix}Cl\\ Cl\end{matrix} \qquad (22)$$

и не явление диссоциации

$$ClCH=CHHgCl \rightleftarrows C_2H_2+HgCl_2, \qquad (23)$$

$$(ClCH=CH)_2Hg \rightleftarrows 2C_2H_2+HgCl_2. \qquad (24)$$

Эти две возможности объяснения двойственности реакций отвергаются фактом существования устойчивых, не переходящих друг в друга (без облучения) цис-трансизомеров в рассматриваемом ряду соединений. Объяснение двойственности реагирования хлорвинильных производных ртути должно опираться, таким образом, на единую структурную формулу каждого из этих веществ.

Чтобы правильно оценить причины особых свойств квазикомплексных соединений, следует обратить внимание, кроме характерной легко-

сти элиминации ацетилена, также на легкость реакций обмена атома ртути. Эти реакции обмена проходят не только значительно легче, чем в случае обычных жирных ртутноорганических соединений, но и чем в случае дифенилртути, вступающей в реакции обмена легче, чем жирные соединения ртути. Это ясно из описанных ниже исключительно легко протекающих реакций обмена с $TlCl_3$, $Pb(OCOCH_3)_4$, реакции с $SnCl_2$ и других, что, очевидно, должно зависеть от характера связи ртути с углеродом, приближающегося к ионному.

Отношения, совершенно аналогичные описанным выше, имеют место и в ряду других полученных нами хлорвинильных производных металлов.

При насыщении пятихлористой сурьмы (лучше в присутствии сулемы) при 150° ацетиленом до прекращения поглощения и при перегонке продукта в вакууме (165—170° при 3 мм) переходит смесь двух кристаллических соединений, которые могут быть разделены кристаллизацией из спирта и затем из легкого петролейного эфира [12].

При этом более высокоплавкое (т. пл. 93—94°) и менее растворимое соединение состава $(C_2H_2)_3SbCl_5$, выход которого достигает 80%, получается в значительно большем количестве. Второе соединение имеет тот же состав и является геометрическим изомером первого; оно плавится при температуре 61—62°. Выход его достигает 3—4%. Легко показать, что оба вещества являются три-β-хлорвинилдихлорстибинами; с иодом они количественно отщепляют 1.2-иодхлорэтилен, с бромом — 1.2-хлорбромэтилен и в этом отношении совершенно подобны своим ртутным аналогам. Восстановлением обоих продуктов гидросульфитом натрия получаются два стереоизомерных три-β-хлорвинилстибина. Следующая таблица дает сопоставление свойств этих четырех аддуктов хлористой сурьмы и ацетилена:

$(ClCH{=}CH)_3SbCl_2$	$(ClCH{=}CH)_3SbCl_2$
α-форма (транс)	β-форма (цис)
т. пл. = 93—94°	т. пл. = 61—62°
$d_{4°}^{100°} = 1.7838$	$d_{4°}^{100°} = 1.7838$
σ 100° = 35.95 эрг/см²	σ 100° = 37.49 эрг/см²
Найдено $P = 517.6$	Найдено $P = 523$

Вычислено $P = 513.7$

$$\begin{array}{l} ClCH=CH\diagdown \\ ClCH=CH-Sb \\ ClCH=CH\diagup \end{array}$$

α-форма (транс)
т. пл. $=48$—$49°$
$d_{4°}^{100°}=1.7021$
$\sigma^{100°}=35.10$ эрг/см²
Найдено $P=437.8$

$$\begin{array}{l} ClCH=CH\diagdown \\ ClCH=CH-Sb \\ ClCH=CH\diagup \end{array}$$

β-форма (цис)
т. к. $=121-122°$ (3 мм)
$d_{4°}^{15°}=1.8070$
$d_{4°}^{100°}=1.7181$
$\sigma^{100°}=35.93$ эрг/см²
Найдено $P=436.3$

Вычислено $P=429.9$

Расчет велся по атомным парахорам Сегдена, принимая, как это он делает для $SbCl_5$, декремент двух сингулентных связей. Определение парахора выполнила Н. К. Гипп.

Вместе с тем все четыре вещества подобны своим ртутным аналогам и в другом отношении: при воздействии агентов, связывающих галогенид сурьмы, они гладко отщепляют ацетилен. Так, все они при действии теплой щелочи количественно выделяют ацетилен:

$$(ClCH=CH)_3SbCl_2+NaOH\rightarrow 3C_2H_2+NaCl+Na_2H_2Sb_2O_7,$$
$$(ClCH=CH)_3Sb+NaOH\rightarrow 3C_2H_2+NaCl+NaSbO_2.$$

Еще легче вытеснение ацетилена происходит при аналогичной реакции с сернистым натрием, разлагающим в водноспиртовом растворе обоих типов соединения уже на холоду и очень быстро. Может быть, еще более наглядно демонстрируется квазикомплексный характер описываемых веществ реакцией их с такими органическими сульфидами, как тиомочевина, при нагревании со спиртовым раствором которых также происходит вытеснение ацетилена. Спиртовый раствор азотнокислого серебра реагирует с $(ClCH=CH)_3Sb$, образуя ацетиленид серебра состава $Ag_2C_2AgNO_3$ и в небольшом количестве комплексное соединение $[(ClCH=CH)_3Sb]_2\cdot AgNO_3$.

Таким образом, и эти хлорвинильные металлоорганические производные сурьмы имеют ярко выраженные черты химического поведения комплексных соединений. Еще легче происходит распад соответствующих первичных и вторичных соединений. При неполном насыщении пятихлористой сурьмы ацетиленом при комнатной температуре образуется

продукт состава $(C_2H_2)_2\,SbCl_5$, который плавится при 52°, имеет $d_{4°}^{100°}$ = 1.9329, $\sigma^{100°}$ = 36.77 эрг/см², $P_{найд}$ = 447.24, $P_{выч}$ = 446.70. Это вещество является ди-β-хлорвинилтрихлорстибином неизвестной пока конфигурации, хотя можно предполагать, что оно относится к трансряду. Действием металлической сурьмы в соответствии с уравнениями

$$3\,(ClCH{=}CH)_2\,SbCl_3 + 2Sb \rightarrow 3\,(ClCH{=}CH)_2SbCl + 2SbCl_3,$$
$$(ClCH{=}CH)_2\,SbCl + SbCl_3 \rightarrow 2\,ClCH{=}CHSbCl_2$$

это соединение можно превратить в смесь хлорвинилдихлорстибина [т. к. 102—105° (1.5 мм), $d_{4°}^{20°}$ = 2.1486] и дихлорвинилхлорстибина (т. к. 113—115° при 2 мм), в которых ацетилен связан настолько непрочно, что при перегонке в вакууме (1—2 мм) постепенно отщепляется; отщепление ацетилена происходит и при хранении веществ. Продукт низшей степени насыщения ацетиленом $C_2H_2SbCl_5$ описан Бертло [12a], а затем Ньюлендом [12b] как комплексное соединение, и нами пока не исследовался. Интересный распад претерпевает при пиролизе (200—250°) более высокоплавкий изомер три-β-хлорвинилдихлорстибина. Этот распад совмещает характерные черты распада металлоорганических соединений с отщеплением свободного радикала и черты распада комплекса — распад на слагающие его, как бы автономные, самостоятельные молекулы. Органическая часть этого вещества отходит на одну треть в виде продукта спаривания двух хлорвинильных радикалов — 1.4-дихлорбутадиена-1.3, на половину в виде ацетилена и на одну шестую в виде дихлорэтилена

$$2\,(ClCH{=}CH)_3\,SbCl_2 \rightarrow ClCH{=}CH-CH{=}CHCl + 3C_2H_2 + 2SbCl_3 + ClCH{=}CHCl.$$

В ряде реакций, как уже сказано, оба три-β-хлорвинилстибиндихлорида и оба три-β-хлорвинилстибина передают свои β-хлорвинильные радикалы без отщепления ацетилена (реакции с галоидами). К этому ряду реакций нужно присоединить реакции обмена три-β-хлорвинилстибинов с сулемой

$$\left(\begin{matrix}Cl\\H\end{matrix}\!\!>C{=}C\!<\!\begin{matrix}H\\ \end{matrix}\right)_3 Sb + 3HgCl_2 \rightarrow 3\,\begin{matrix}Cl\\H\end{matrix}\!\!>C{=}C\!<\!\begin{matrix}H\\HgCl\end{matrix} + SbCl_3,$$

$$\left(\begin{matrix}H\\Cl\end{matrix}\!\!>C{=}C\!<\!\begin{matrix}H\\ \end{matrix}\right)_3 Sb + 3HgCl_2 \rightarrow 3\,\begin{matrix}H\\Cl\end{matrix}\!\!>C{=}C\!<\!\begin{matrix}H\\HgCl\end{matrix} + SbCl_3.$$

Реакция более высокоплавкого три-β-хлорвинилстибина дает при этом транс-β-хлорвинилмеркурхлорид — «комплекс Бигинелли», а низкоплавкий изомер образует цисизомер.

Аналогичными свойствами квазикомплексных соединений обладают и полученные в нашей лаборатории хлорвинильные производные олова [13] и свинца [15].

Два стереоизомерных хлористых ди-β-хлорвинилолова [13] получаются при взаимодействии обоих изомеров ди-β-хлорвинилртути с хлористым оловом в абс. спирте по реакции Несмеянова и Кочешкова [14], которая течет в этом случае исключительно легко уже на холоду и завершается при 45—50°:

$$(ClCH=CH)_2Hg + SnCl_2 \rightarrow (ClCH=CH)_2SnCl_2 + Hg.$$

Жидкая ди-β-хлорвинилртуть (цисизомер) дает жидкий хлорид ди-β-хлорвинилолова, перегоняющийся при 3 мм при 100—102°, имеющий $d_{4^\circ}^{20^\circ} = 1.7494$ и $n_D^{20^\circ} = 1.5675$, а $MR = 58.40$. Транс-ди-β-хлорвинилртуть образует с $SnCl_2$ кристаллический стереоизомер хлористого ди-β-хлорвинилолова — вещество с т. пл. 77.5—78.5°, разлагающееся лишь при 135°. В соответствии с высокой точкой плавления это вещество является трансизомером, что подтверждается его реакцией обмена с сулемой, приводящей количественно к транс-β-хлорвинилмеркурхлориду

$$(ClCH=CH)_2SnCl_2 + 2HgCl_2 \rightarrow 2ClCH=CHHgCl + SnCl_4,$$

тогда как жидкий стереоизомер с сулемой дает только низкоплавкий (цис) стереоизомер хлористой хлорвинилртути.

Оловоорганические хлорвинильные производные довольно прочны, особенно цисизомер, разлагающийся по квазикомплексному типу лишь при действии 50% щелочи (на холоду); транссоединение разлагается уже 15% едким натром.

Хлорвинильные производные свинца получены нами и А. К. Кочетковым [15] из транс-ди-β-хлорвинилртути посредством реакции К. А. Кочешкова и М. М. Надь [16]

$$2(ClCH=CH)_2Hg + Pb(OOCCH_3)_4 \rightarrow (ClCH=CH)_2Pb(OCOCH_3)_2 + \\ + ClCH=CHHgOOCCH_3.$$

Ацетат дихлорвинилсвинца представляет собою хорошо растворимый в хлороформе кристаллический продукт, при нагревании разлагающийся (не плавясь) с выделением ацетилена. С ацетатом аммония образует молекулярное соединение $(ClCH=CH)_2Pb(OOCCH_3)_2 \cdot NH_4OOCCH_3$, разлагающееся при 177—178°. Спиртовый раствор последнего с эквивалент-

ным количеством хлороводорода дает дихлорид $(ClCH=CH)_2PbCl_2$ (т. разл. 163—167°). Получены также нерастворимый, неплавкий сульфат и бензоат с т. пл. 204°. Судя по реакции с металлической ртутью

$$(ClCH=CH)_2PbCl_2+Hg \rightarrow ClCH=CHHgCl$$

и реакции с хлористым оловом

$$(ClCH=CH)_2PbCl_2+SnCl_2 \rightarrow (ClCH=CH)_2SnCl_2,$$

при которых образуются только трансмодификации, рассматриваемое свинцовоорганическое соединение, вероятно, также относится к трансряду.

Все эти хлорвинильные соединения свинца — также типичные квазикомплексные соединения: при действии водного раствора иодистого калия, при восстановлении гидразингидратом, при фенилировании бромистым фенилмагнием они элиминируют ацетилен. В последней реакции свинец выделяется в виде тетрафенилсвинца. При термическом распаде и действии минеральных кислот также (но не количественно) выделяют ацетилен.

При сливании эфирных растворов транс-ди-β-хлорвинилртути и треххлористого таллия осаждается смесь хлорида дихлорвинилталлия и хлорвинилмеркурхлорида

$$2\,(ClCH=CH)_2Hg+TlCl_3 \rightarrow (ClCH=CH)_2TlCl+2ClCH=CHHgCl,$$

из которой последний отмывается бензолом, а первый перекристаллизовывается из водного спирта. Это вещество плавится при 160—162° с разл. Оно выделяет ацетилен при действии щелочи или реактива Илосвая, а также при действии разбавленной соляной кислоты [17].

Поразительна легкость и быстрота протекания всех разнообразных описанных реакций обмена металла в хлорвинильных соединениях, далеко превосходящая все, что известно не только для жирных, но и для ароматических металлоорганических соединений.

Может быть, наиболее неожиданны квазикомплексные свойства у хлорвинильных производных многовалентного иода

$$ClCH=CHJCl_2,\ [18]\ (ClCH=CH)_2JCl \text{ и } (ClCH=CH)(C_6H_5)JCl\ [19],$$

из которых первое может быть получено или хлорированием хлориодэтилена [20], или присоединением треххлористого иода к ацетилену [21], или, наконец, действием треххлористого иода на транс-ди-β-хлорвинилртуть,

второе — действием первого на транс-ди-β-хлорвинилртуть, а третье — реакцией транс-ди-β-хлорвинилртути с $C_6H_5JCl_2$ [19].

Все три соединения при нагревании, действии пиридина, KCl и, лучше, CsCl, выделяют (но далеко не количественно) ацетилен, в соответствии со схемой

$$ClCH=CHJCl_2+CsCl \rightarrow C_2H_2+CsJCl_4.$$

Значительно ярче эти свойства выражены, как это показывает исследование Фрейдлиной и Брайниной [22], у β-иодвинилиодидхлорида, который под действием пиридина уже на холоду и почти количественно отщепляет ацетилен, а при арилировании фенилтрихлорстаннаном также элиминирует ацетилен и образует продукт фенилирования трехгалоидного иода — соль дифенилиодония

$$JCH=CHJCl_2+2C_6H_5SnCl_3 \rightarrow (C_6H_5)_2JCl+2SnX_4+C_2H_2,$$

где $X=Cl, J$.

Этот факт проливает некоторый свет на природу явления. В отщеплении галоидного соединения элемента ($HgCl_2$, JCl_3 и т. д.) от квазикомплексного соединения не играет, очевидно, никакой роли возникающая (или, предположительно, промежуточно могущая возникнуть) связь элемента и галоида, поскольку нельзя предположить, что такую роль играет связь $J-JCl_2$. Что это отщепление тем более не обусловлено также предсуществованием какой-либо химической связи между металлом и галоидом в молекуле квазикомплексного соединения, например:

$$\begin{matrix} H \\ Cl \end{matrix} > C = C < \begin{matrix} H \\ HgCl \end{matrix} \quad (Cl \cdots Hg),$$

следует с несомненностью из наличия квазикомплексных свойств у цис- и у транссоединений с удаленным положением атомов металла и галоида.

КВАЗИКОМПЛЕКСНЫЕ АДДУКТЫ ОЛЕФИНОВ И ОКИСИ УГЛЕРОДА

Аналогичные свойства квазикомплексных соединений принадлежат также открытым Гофманом и Зандом [23] веществам, получаемым присоединением солей ртути к олефинам. Этого типа производные известны в настоящее время для широкого ряда соединений, содержащих двойные C = C-связи. В зависимости от растворителя, в котором происходит присоединение

соли ртути к олефину и который всегда принимает участие в реакции, образуются такие типы соединений (например, этилена):

$$CH_2{=}CH_2+HgX_2+H_2O \rightarrow HOCH_2CH_2HgX+HX, \quad [23]$$

$$2CH_2{=}CH_2+2HgX_2+H_2O \rightarrow O(CH_2CH_2HgX)_2+2HX, \quad [24]$$

$$CH_2{=}CH_2+HgX_2+AlkOH \rightarrow AlkOCH_2CH_2HgX+HX, \quad [25]$$

$$CH_2{=}CH_2+Hg(OCOR)_2 \rightarrow RCOOCH_2-CH_2HgOOCR, \quad [26]$$

$$CH_2{=}CH_2+HgX_2+R_2NH \rightarrow R_2NCH_2-CH_2HgX+HX, * \quad [27]$$

$$CO+HgX_2+AlkOH \rightarrow AlkO-\overset{\displaystyle O}{\overset{\|}{C}}-HgX+HX, \quad [28]$$

$$CO+HgX_2+R_2NH \rightarrow R_2N-\underset{\displaystyle O}{\underset{\|}{C}}-HgX+R_2NH\cdot HX\ {**}. \quad [29]$$

Всем этим веществам, несомненно, принадлежит приданное строение, так как они гладко замещают ртуть на водород при восстановлении амальгамой натрия в водной среде и образуют соответствующий спирт, эфир или третичный амин, а также замещают ртуть на атом галоида при реакциях галоидирования. Нетрудно доказать строение также и характерной реакцией на функциональную группу, например, ацилированием алканолмеркурхлорида посредством хлористого бензоила

$$HOCH_2CH_2HgCl+C_6H_5COCl \rightarrow C_6H_5COOCH_2CH_2HgCl \quad [30]$$

или фенилизоцианата

$$HOCH_2CH_2HgCl+C_6H_5NCO \rightarrow C_6H_5NHCOOCH_2CH_2HgCl \quad [31]$$

или путем солеобразования за счет третичного азота аминоалкилмеркурсоли

$$R_2NCH_2CH_2HgCl+HCl \rightarrow [R_2HNCH_2CH_2HgCl]\,Cl. \quad [27]$$

Связанная с углеродом галогенмеркур-группа ClHg ведет себя в ряде реакций соответственно обычным ртутноорганическим соединениям, подобным метилмеркурхлориду CH_3HgCl; например, заменяет галоид на арил при действии оловоорганических соединений

$$HOCH_2CH_2HgCl+C_6H_5SnCl_3+NaOH \rightarrow HOCH_2CH_2HgC_6H_5+Na_2SnO_3+$$
$$+NaCl+H_2O.$$

---

* Этот тип веществ впервые получен Р. Х. Фрейдлиной и Н. С. Кочетковой.

** Этого рода соединения получены Р. Х. Фрейдлиной и Э. И. Кан.

При действии диазометана также происходит в известной мере аналогичная реакция замены галоида на хлорметильный радикал

$$HOCH_2CH_2HgCl + CH_2N_2 \rightarrow HOCH_2CH_2HgCH_2Cl + N_2. \qquad [32]$$

Однако при взаимодействии с уже упомянутым ранее рядом реагентов, связывающих соль ртути в комплекс (иодистым натрием, цианистым калием, тиосульфатом натрия и т. д.), происходит на холоду во всех без исключения типах веществ столь же легкая, быстрая и количественная элиминация олефина, как это было описано для ацетилена. Алканолмеркурсоли и их эфиры также легко и количественно распадаются с отщеплением олефинов и под действием даже разбавленных минеральных кислот, но на этот раз это свойство не разделяется β-аминоалкилмеркурсолями. Так, N-β-пиперидил-этилмеркурхлорид устойчив к действию крепкой соляной, серной и хлорной кислот (но не к действию HJ) даже при длительном нагревании, что далеко превосходит устойчивость обычных ртутноорганических соединений. Это подчеркивает, что механизм распада алканолмеркурсолей при действии агентов комплексообразователей и кислот разный и что действие водородного иона направлено не на ртутный конец молекулы, как это очевидно для действия ионов иода, циана и т. д., а на атом кислорода и соотв. азота.

Как и для продуктов, происшедших путем присоединения ртутных солей к ацетилену, совершенно исключительная легкость элиминации олефина из алканолртутных солей (и их эфиров) дала повод рассматривать эти вещества как комплексные соединения типа

$$C_2H_4 \cdot Hg\begin{smallmatrix}\diagup OH \\ \diagdown Cl\end{smallmatrix}, \qquad C_2H_4 \cdot Hg\begin{smallmatrix}\diagup OCH_3 \\ \diagdown Cl\end{smallmatrix}$$

и т. п., а также вызвала дискуссию об их алканольном (соотв. эфирном) или комплексном строении или таутомерии [33].

Следует отметить, что только в одном типе соединений известен аддукт сулемы по двойной связи — продукт присоединения к кумарину

$$C_6H_4\begin{smallmatrix}CHCl-CH(HgCl) \\ | \qquad\quad | \\ O \;\;-\;\; CO\end{smallmatrix}.$$

Во всех остальных случаях галоидная ртуть присоединяется к олефинам лишь при подщелочении и так, что по одну сторону двойной связи

встает остаток HgCl, а по другую гидроксил (в спиртовой среде алкоксил). Все попытки заменить этот гидроксил на хлор (действием, например, хлористого тионила и т. п.) кончаются отщеплением олефина.

В заключение интересно сопоставить рассмотренные типы ртутных аддуктов по их устойчивости

$$(ClCH{=}CH)_2Hg > \underset{\text{не существует}}{(JCH{=}CH)_2Hg}$$

$$R_2NCH_2CH_2HgCl > HOCH_2CH_2HgCl \gg \underset{\text{не существует}}{ClCH_2CH_2HgCl}.$$

Ряд построен в порядке возрастающей способности R_2N, HO, Cl нести анионный заряд. Известны многочисленные аддукты олефинов и галогенидов

Zn[34], Al[35], Sb[36], Fe[37], Pd[38], Pt,[39] Cu[40],

которые считаются комплексными соединениями, но строение которых исследовано слишком мало, чтобы можно было уверенно решить, являются ли они действительно комплексными или квазикомплексными β-галоидозамещенными металлоорганическими соединениями.

ПРИЧИНА КВАЗИКОМПЛЕКСНЫХ СВОЙСТВ

Отношения и свойства широкого ряда продуктов присоединения металлических солей к ацетилену и олефинам так сходны, что причины их особых свойств, делающих их столь похожими на комплексные соединения, должны быть общими. Ясно, что эти причины не могут заключаться ни в подвижном равновесии аддукта с продуктами его диссоциации, например:

$$ClCH{=}CHHgCl \rightleftarrows C_2H_2 + HgCl_2,$$

ни в новом виде таутомерии — обратимом изомерном превращении истинного хлорвинильного металлоорганического соединения в истинное комплексное

$$ClCH{=}CHHgCl \leftrightarrows C_2H_2 \cdot HgCl_2,$$

прежде всего потому, что наличие того или другого явления исключило бы возможность устойчивого существования цис-трансизомеров аддуктов. Поскольку строение β-замещенных металлоорганических веществ безукоризненно доказано во всех разобранных нами случаях, нужно

объяснить особенности квазикомплексного поведения, исходя из структур типа

$$\begin{matrix}Cl \\ H\end{matrix}\!\!>C=C<\!\!\begin{matrix}H \\ HgCl\end{matrix},$$

$$HO-CH_2-CH_2HgCl \text{ и т. п.}$$

При этом следует принять во внимание значительно бо́льшую «подвижность» ртути, близость ее к ионному состоянию, чем во всех обычных ртутноорганических соединениях, что выражается не только в легкости ее отщепления комплексообразователями (J', CN', $P(C_6H_5)_3$ и т. п.), но и в исключительной легкости реакций обмена, превосходящей легкость обмена огромного большинства ртутноорганических соединений. Это же относится к подвижности гидроксильной группы этанолмеркурсолей, также проявляющейся не только в описанных выше реакциях элиминации олефина, но и в реакциях дегидратации под влиянием, например, ацетамида (Фрейдлина, Кочетков)

$$2ClHgCH_2-CH_2OH \rightarrow (ClHgCH_2CH_2)_2O$$

и к подвижности хлора в хлорвинильных производных ртути. «Предсуществование» молекулы ацетилена или олефина в квазикомплексных соединениях и, так сказать, «федеративная» структура последних, подвижность — приближение к ионному состоянию металла и анионной части аддукта — наиболее естественно могут быть объяснены эффектом гиперконъюгации, выражаемым схемами такого типа:

$$\begin{matrix}Cl \\ H\end{matrix}\!\!>C=C<\!\!\begin{matrix}H \\ HgCl\end{matrix}$$

$$\begin{matrix}Cl \\ H\end{matrix}\!\!>C=C<\!\!\begin{matrix}Hg \\ H\end{matrix}\qquad \begin{matrix} \\ H\end{matrix}\!\!>C=C<\!\!\begin{matrix}Cl \\ H\end{matrix}$$

$$HO-CH_2-CH_2-HgCl \qquad\qquad R_2N-\underset{\underset{O}{\|}}{C}-HgBr$$

1,2 — сдвиг 1,1 — сдвиг

При этом, вероятно, в основе такого сдвига, совершающегося до конца под влиянием атакующего реагента, лежит более или менее выраженный постоянный структурный — мезомерный (≡резонансный) эффект, который можно также выразить в символике Паулинга так:

$$\begin{matrix}Cl\\ H\end{matrix}\!\!>C=C\!\!<\!\!\begin{matrix}H\\ HgCl\end{matrix} \longleftrightarrow \begin{matrix}Cl'\\ H\end{matrix}\!\!/C\equiv C/\!\!\begin{matrix}H\\ \overset{+}{Hg}\,Cl\end{matrix}$$

$$HO-CH_2-CH_2-HgCl \longleftrightarrow HO'CH_2=CH_2\,\overset{+}{Hg}Cl$$

$$R_2N-\underset{\underset{O}{\|}}{C}-HgBr \longleftrightarrow R_2N'\,\underset{\underset{O}{|||}}{C}\overset{+}{Hg}Br$$

Наличие мезомерного сдвига сравнительно с хлорвинильной структурой ясно для обоих стереоизомеров β-хлорвинилмеркурхлорида на основании сравнения величин их дипольных моментов с рассчитанными для хлорвинильных структур.

Дипольный момент трансхлорвинилмеркурхлорида (в бензоле) 1.34 *D*. Рассчитанный из дипольных моментов хлористого винила и связи Hg—Cl (взятой из хлористой бензилртути), он был бы 1.64 *D*. Для цисизомера дипольный момент найденный 2.76 *D*, а рассчитанный 4.08 *D*. Налицо, таким образом, значительное отклонение от аддитивности, которое может быть объяснено перекрыванием в обратном направлении слагающей ионной структуры типа

$$\begin{matrix}Cl'\\ H\end{matrix}\!\!/C\equiv C/\!\!\begin{matrix}H\\ \overset{+}{Hg}\,Cl\end{matrix}$$

с весом 20—30%.

Особенностью типа резонанса в разбираемых квазикомплексных соединениях является выделение в одной из канонических структур ничем не связанных частей молекулы, способных к самостоятельному существованию

$$\begin{matrix}Cl'\\ H\end{matrix}\!\!/C\equiv C/\!\!\begin{matrix}H\\ \overset{+}{Hg}\,Cl\end{matrix} \qquad HO'CH_2=CH_2\overset{+}{Hg}\,Cl,$$

что и определяет «федеративный» характер структуры и намечает «трещины», по которым происходит распад.

Мы полагаем, что эта постоянная мезомерная часть всего эффекта иногда может быть и невелика и дополняется динамическим электромерным сдвигом электронов, осуществляемым лишь в момент реакции, и что

поэтому, может быть, правильнее применение символики Робинзона — Ингольда, суммирующей оба эффекта:

$$\begin{matrix} Cl & & H \\ & C=C & \\ H & & HgCl \end{matrix}$$

Об этом можно судить на том основании, что и среди чисто органических веществ такого типа, в котором (на основании физических свойств) вряд ли можно ожидать резко выраженного явления мезомерии этого рода, известны явления реакций и распада, подобные реакциям и распаду квазикомплексных соединений. Таковы примеры хлоральгидрата (и подобных гидратов), реагирующего как альдегид

$$CCl_3-\overset{H}{C}\begin{matrix} O-H \\ O-H \end{matrix} \quad \text{или} \quad CCl_3-\overset{H}{\underset{\| O}{C}} \quad \overset{H}{O} \quad H^+$$

«дифосгена», реагирующего в ряде случаев как две молекулы фосгена

$$Cl-\underset{\| O}{C}-O-\underset{Cl}{\overset{Cl}{C}}-Cl \quad \text{или} \quad Cl-\underset{\| O}{C}^{+} \quad O=C\begin{matrix} Cl \\ Cl \end{matrix} \quad Cl^{-}$$

хлоругольно-этилового эфира

$$Cl-\underset{\| O}{C}-O-CH_2-CH_2-H \quad \text{или} \quad Cl^{-} \quad \underset{\| O}{C}=O \quad CH_2=CH_2H^+$$

легко распадающегося при воздействии диметилцинка с образованием этилена, наконец, открытого Несмеяновым и Кан фтористого формила [41], легко распадающегося на CO и HF

$$H \frown C(=O) \frown F \quad \text{или} \quad H^{+}\; C(=O{:})\; F' \quad \text{или} \quad H^{+}\; C{\equiv}O\; F'$$

Все соединения этой области органических веществ имеют черты «федеративной» структуры и в этом отношении представляют полную аналогию с квазикомплексными соединениями; последние отделяются лишь искусственно, по признаку содержания в них в качестве федеративной части молекулы металлической соли.

КВАЗИКОМПЛЕКСНЫЕ ВЕЩЕСТВА, ПЕРЕНОС РЕАКЦИОННОГО ЦЕНТРА ВСЛЕДСТВИЕ КОНЪЮГАЦИИ И ГИПЕРКОНЪЮГАЦИИ И ТАУТОМЕРИЯ

Академик А. Н. Несмеянов и И. Ф. Луценко

С другой стороны, ртутноорганические квазикомплексные вещества хлорвинильного или этанольного типа по особенностям своего строения тесно примыкают к изученным в последнее время Несмеяновым и Луценко [42] хлормеркуральдегидам и кетонам, что видно из следующего сопоставления

$$ClHg - CH = CH - Cl$$

$$ClHg - CH_2 - CH_2 - OH$$

$$ClHg - CH_2 - C(H) = O$$

Так как эта новая аналогия проливает свет на некоторые черты поведения квазикомплексных соединений, то необходимо несколько подробнее осветить свойства галогенмеркурацетальдегида и ацетона.

Они получены реакцией присоединения ацетата ртути в водной среде к простым или сложным виниловым (соотв. пропениловым) эфирам с последующим осаждением хлористым натрием

$$(CH_3COO)_2Hg + CH_2=CH-OC_4H_9 + H_2O \rightarrow$$
$$\rightarrow CH_3COOHgCH_2-C\begin{smallmatrix}\nearrow O\\ \searrow H\end{smallmatrix} + CH_3COOH + C_4H_9OH \xrightarrow{Cl'}$$
$$\rightarrow ClHgCH_2-C\begin{smallmatrix}\nearrow O\\ \searrow H\end{smallmatrix}$$

Строение хлормеркурацетальдегида доказано гладким восстановлением изопропилатом алюминия в этанолмеркурхлорид и почти количественным превращением действием брома в бромацетальдегид. Особенностью этого вещества, разделяемой им с квазикомплексными ртутноорганическими соединениями, является высокая лабильность ртути. Вещество не дает обычных альдегидных реакций: при действии гидроксиламина, фенилгидразина, гидразина — выделяется металлическая ртуть. Сероводород осаждает сернистую ртуть (щелочь, однако, не дает желтого осадка окиси ртути), при подкислении выделяется ацетальдегид. Аналогично ведет себя и хлормеркурацетон.

Решающей для суждения об известной аналогии с квазикомплексными соединениями в смысле характера особенностей резонанса молекулы является реакция ацилирования. При действии хлористого ацетила в качестве единственного продукта реакции получен винилацетат, при действии хлористого бензоила — винилбензоат, что можно объяснить только написанным выше эффектом сдвига электронов и сосредоточением отрицательного заряда на кислороде, как бы энольным сдвигом электронов.

$$ClHg \frown CH_2 \frown C(H)=O + CH_3-C(=O)-Cl \longrightarrow Cl_2Hg + CH_2=C(H)-O-C(=O)-CH_3$$

Как видно, этот «энольный сдвиг электронов» совершенно аналогичен рассмотренным выше для хлорвинилмеркурхлорида и хлормеркурэтанола

$$ClHg \frown CH = CH \frown Cl$$
$$ClHg \frown CH_2 - CH_2 \frown OH$$

с той разницей, что в первом случае взаимодействуют σ и π-связь, в случае же квазикомплексных соединений — две σ-связи. Случай хлормеркурацетальдегида является таким образом промежуточным между случаем классического сопряжения (конъюгации) двух двойных (π)-связей, например, $C{=}C - C{=}O$, и здесь разбираемым сопряжением простых (σ)-связей, обусловливающим особенности квазикомплексной структуры. Легко видеть, что данный эффект есть не что иное, как так называемый эффект гиперконъюгации [43], вскрытый Натаном и Бэкером [44] в поведении метильной группы и, в частности, на примере уксусного альдегида, выражающийся схемой

$$H_3{\equiv}C - \underset{\displaystyle H}{C} = O$$

Однако в силу особенности связи $Hg - C$, значительно более способной к электронному сдвигу в направлении стрелки, в случае ртутноорганических соединений этот эффект выражен гораздо более ярко. Именно в силу этой резкости эффекта для ртутных производных он так ярко заметен не только на примере ацетальдегидного деривата, но и на примере хлормеркурэтанола и хлорвинильных соединений металлов, тогда как ни в этаноле, ни в хлористом виниле явление никак не проявляется.

Для характеристики явления гиперконъюгации интересно отметить, что, например,

$$\left(\begin{matrix} Cl & & H \\ & C = C & \\ H & & \end{matrix}\right)_2 Hg$$

в сопряжении участвует не только $Hg - C$ σ-связь, но ее резко выраженное участие одновременно вовлекает и $H - C$ σ-связь в гораздо большей степени, чем, например, это имеет место в хлористом виниле. Это проявляется в том, что водородные атомы получают значительную подвижность. При действии не только сильно щелочных агентов (алкоголят натрия) и жирных аминов (этилендиамин), но даже

уксуснокислого натрия и фтальимида калия происходит отщепление хлороводорода и образуется нерастворимый ацетиленид ртути, при подкислении отдающий весь углерод в виде ацетилена (Р. Х. Фрейдлина и А. К. Кочетков).

Если бы связь ртути с β-углеродным атомом в хлормеркурацетальдегиде не была установлена однозначно, поведение его в реакции ацилирования было бы, несомненно, истолковано как поведение таутомерного оксоэнольного вещества

$$\mathrm{Cl-Hg-CH_2-\underset{\diagdown H}{C}=O \rightleftarrows CH_2=C\langle^{O-HgCl}_{H}}$$

так, как это, например, и сделано в свое время В. Хюккелем для случая натриевых производных кетоэнолов, например

$$\mathrm{CH_3-\underset{\underset{\underset{Na}{|}}{\overset{|}{O}}}{C}=CH-COOC_2H_5 \rightleftarrows CH_3-\underset{\overset{\|}{O}}{C}-\underset{\overset{|}{Na}}{CH}-COO\,C_2H_5}.$$

До этого установилась точка зрения об энольном строении натриевых производных кетоэнолов. Ясно, что О-ацилирование С-металлопроизводного оксосоединения в найденных нами случаях и широко известное С-ацилирование (или алкилирование) О-металлических производных оксосоединений — явления сходно противоположные, и наличие твердо установленного одного из них укрепляет возможность аналогичной трактовки другого.

В настоящее время наиболее распространена иная точка зрения на способность двоякого реагирования натриевых производных кетоэнольных ссединений, а именно представление ионной связи натрия с мезомерным анионом кетоэнола.

Между тем эта точка зрения очевидным образом не применима к двойственному поведению хлормеркурацетальдегида, который хотя и содержит лабильную, но отнюдь не ионную ртуть, связанную в альдегидной форме молекулы, что не мешает последней реагировать по энольному типу. Эта точка зрения не применима и к известным случаям С-ацилирования медного производного ацетоуксусного эфира, заведомо построенного по О-типу, так как оно имеет циклическую («клешнеобраз-

ную») структуру. Делается очевидным, что ни предположение Хюккеля о таутомерном равновесии C — Na и O — Na форм натриевых производных кетоэнолов, ни современное толкование их двойственного поведения на основе мезомерии аниона не являются необходимыми для объяснения их двойственной реакционной способности. Ясно по аналогии, что структура, например, натрий-ацетоуксусного эфира может подвергаться C-алкилированию или ацилированию на основании того же явления гиперконъюгации

$$CH_3-\underset{\underset{\underset{Na}{|}}{\overset{|}{O}}}{C}=CH-COOC_2H_5$$

создающего отрицательный центр на углеродном атоме. Такое объяснение, распространенное на иные случаи «таутомерии», лучше согласуется с общеизвестным поведением в реакции алкилирования солей разных металлов одного и того же соединения, например, O-метилирования натриевых и N-метилирования серебряных соединений изатина или бензамида, что было бы невозможно, если бы метилировался один и тот же для обеих солей мезомерный анион. Естественно допустить, что ковалентная Na — O связь ближе к ионному типу, чем Ag — O связь, и поэтому явление гиперконъюгации в первой связи сильнее, в результате чего и наведенный заряд на атоме азота в случае Na — O производного больше, чем в случае Ag — O производного

C=O, C—O—Na, N   C=O, C—O—Ag, N

Аналогичные примеры известны и для диадной таутомерии, например, в случае трифенилметилирования натриевых и серебряных солей диалкилфосфористой кислоты, по Арбузову:

$$Na-O-P(OC_2H_5)_2 + (C_6H_5)_3CCl \rightarrow O=P(OC_2H_5)_2C(C_6H_5)_3 + NaCl$$

$$Ag-O-P(OC_2H_5)_2 + (C_6H_5)_3CCl \rightarrow (C_6H_5)_3C-O-P(OC_2H_5)_2 + AgCl$$

серебряная соль дает О-тритиловый эфир диэтилфосфористой кислоты, а натриевая соль — диэтиловый эфир тритилфосфиновой кислоты. Следует, однако, отметить, что в диадной таутомерии такое объяснение связано с предположением о расширении октета и поэтому может иметь ограниченное применение.

Двойственная реакционная способность металлических производных энолов, лактимов и других родственных соединений, таким образом, в известном отношении глубоко родственна по своим причинам двойственной реакционной способности квазикомплексных соединений. На протяжении истории учения о таутомерии реакции металлических производных прототропных таутомерных систем (или систем, в которых подозревались таутомерные отношения) никогда не отграничивались как явление особого рода от поведения самих прототропных таутомерных систем, хотя только для последних реально установлено существование таутомеров и изучено их равновесие. Такой, например, классический и исторически первый подробно изученный случай «таутомерии», как «таутомерия» изатина, по существу является только примером двойственной реакционной способности металлических производных изатина, но не проявлением таутомерии изатина, для которого существует только одна — лактамная — форма.

Если продолжать причислять такого рода явления к явлениям таутомерии, то нужно причислить к этим явлениям и те отношения двойственной реакционной способности квазикомплексных соединений и хлормеркур-

ацетальдегида, которые подробно обсуждены в настоящей статье. Эти отношения отличаются только своей большей ясностью, невозможностью подозревать участие в реакциях еще не открытой таутомерной формы (псевдомерной формы).

Гораздо правильнее считать таутомерией только область явлений, предвиденных Бутлеровым, — область равновесных обратимых изомеризаций.

Тогда отграничивается другая область — явлений двойственной реакционной способности — и область переноса реакционного центра молекулы в результате явлений резонанса и частного случая резонанса — гиперконъюгации. Сюда относятся и химия «таутомерных» металлических производных, и химия квазикомплексных металлоорганических соединений, и такие случаи подозреваемой таутомерии, как «таутомерия» толуола, по Шорыгину, и многочисленные случаи передачи основных свойств одним атомом, донором, другому (в азокрасителях от аминного азота — азоазоту, в нитрозодиалкиланилинах от азота — кислороду нитрозогруппы), и многочисленные явления ориентации замещения, в которых углеродный атом (например, бензольного ядра) приобретает свою нуклеофильную или электрофильную реакционную способность в результате передачи ее от ключевого атома группы — ориентанта.

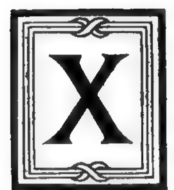
ЛИТЕРАТУРА

1. *А. М. Бутлеров.* Ann Chem. **189**, 76 (1877).
2. *Biginelli.* C. 1898, 1925.
3. *Chapman, Jenkins.* Journ. Chem. Soc. **115**, 849 (1919). *Jenkins.* Journ. Chem. Soc. **119**, 747 (1921).
4. *Фрейдлина, Несмеянов.* ДАН **26**, 59 (1940).
5. *Фрейдлина, Ногина.* Изв. АН ОХН (1947), 99.
6. *Несмеянов, Борисов.* Изв. АН ОХН (1945), 147.
7. *Несмеянов, Фрейдлина, Борисов.* Изв. АН ОХН (1945), 137.
8. *Несмеянов, Борисов, Гуськова.* Изв. АН ОХН (1945), 640.
9. *Китайгородский.* Изв. АН ОХН (1945), 170.
10. *Фрейдлина.* Изв. АН ОХН (1942), 16.
11. *Несмеянов, Фрейдлина.* Изв. АН ОХН (1945), 150.
12. *Несмеянов, Борисов.* Изв. АН ОХН (1945), 251.
12a). *Berthelot, Jungfleisch.* Ann. Chem. Spl. **7**, 252 (1870).
12b). *Nieuwland,* диссертация, 1904, Notre Dame University.
13. *Несмеянов, Борисов, Абрамова.* Изв. АН ОХН (1946), 647.

14. *Несмеянов, Кочешков.* Ber. **63**, 2496 (1930), ЖРФХО **62**, 1795 (1930).
15. *Несмеянов, Фрейдлина, А. Кочетков.* Изв. АН ОХН (1947).
16. *Надь, Кочешков.* ЖОХ **12**, 409 (1942).
17. *А. Кочетков, Фрейдлина, Несмеянов.* Изв. АН ОХН (1947).
18. *Несмеянов, Фрейдлина.* Изв. АН ОХН (1943), 152.
19. *Фрейдлина, Брайнина, Несмеянов.* Изв. АН ОХН (1945), 647.
20. *Thiele, Haak.* Ann. Chem. **369**, 135 (1909).
21. *Howel, Noyes.* J. Amer. Chem. Soc. **42**, 991 (1920).
22. *Брайнина*, диссертация, Институт Органической Химии АН СССР, 1947.
23. *Hoffmann, Sand.* Ber. **33**, 1340, 1359, 2692 (1900).
24. *Hoffmann, Sand.* Ber. **34**, 2907 (1901).
25. *Schoeller, Schrauth, Struensee.* Ber. **43**, 695 (1910) **44**, 1048 (1911).
26. *Hügel, Hibou.* Chim. et Ind. **21**, 296 (1929).
27. *Фрейдлина, Кочеткова.* Изв. АН ОХН (1945), 128.
28. *Schoeller, Schrauth, Essers.* Ber. **46**, 2864 (1913).
29. *Фрейдлина, Кан, Несмеянов.* Изв. АН ОХН (1947).
30. *Sand* **34**, 1385 (1901).
31. *Несмеянов, Фрейдлина.* ЖОХ **7**, 43 (1937).
32. *Фрейдлина, Несмеянов, Токарева.* ЖОХ **7**, 262 (1937),
33. Обзор литературы см. *Несмеянов* и *Фрейдлина.* ЖОХ **7**, 43 (1937).
34. *Кондаков,* ЖРФХО **25**, 356 (1893).
35. *Henderson, Gangloff.* J. Am. Chem. Soc. **38**, 1384 (1918).
36. *Rosenheim, Levy.* Ber. **37**, 3662 (1904).
37. См. 36, но также *Manchot.* Haas. Ber. **45**, 3052 (1912).
38. *Rharasch, Seyler, Mayo.* J. Am. Chem. Soc. **60**, 882 (1935).
39. *Zeise. Pogg.* Ann. **21**, 497 (1831); *Birnbaum.* Ann. Chem. **145**, 67 (1868). *Anderson.* Journ. Chem. Soc. 971 (1934). *Pace.* C. 1213 (1930). *Гельман,* ДАН **16**, 359 (1937).
40. *Manchot.* Ann. **387**, 252 (1912).
41. *Несмеянов, Кан.* ЖОХ **4**, 1247 (1934).
42. *Несмеянов, Луценко.* ДАН (1947).
43. *Mulliken, Rieke, Broun.* J. Am. chem. Soc. **63**, 41 (1941). *Wheland.* J. chem. Phys. **2**, 474 (1934).
44. *Nathan, Baker.* J. chem. Soc. 1844 (1935).

Академик
В. М. Родионов

РАБОТЫ А. П. ОРЕХОВА В ОБЛАСТИ ХИМИИ АЛКАЛОИДОВ



Химия алкалоидов принадлежит к числу очень молодых и очень трудных разделов органической химии. Впервые понятие алкалоид было применено к морфину, соединению, выделенному из опия в 1806 г. немецким аптекарем Сертюрнером. В настоящее время понятие алкалоид применяется к встречающимся в природе основным азотсодержащим соединениям и ограничивается обычно веществами растительного происхождения.

Первое сообщение о выделении морфия (или морфина) из опия не произвело большого впечатления, и лишь опубликование в 1817 г. второй работы Сертюрнера с подробным описанием щелочных свойств морфия вызвало огромный интерес к этому новому и оригинальному виду органических соединений.

С 1817 г. началось продолжающееся и до настоящего времени тщательное изучение многих сложных медицинских препаратов («дрог»), а также растительных экстрактов с целью выделения из них действующих начал. Результаты быстро сказались. Особенно здесь надо отметить работы французских ученых Пелетье и Кавенту. В продуктах растительного происхождения (хинная корка, рвотный орех, бобы Игнатия, табак, кофе, чай, бобы какао и т. п.) было открыто много физиологически и фармакологически интересных и ценных препаратов.

Русские химики принимали небольшое участие в этих исследованиях. Можно указать лишь на одну работу замечательного русского ученого А. А. Воскресенского, выделившего в 1842 г. из бобов какао алкалоид теобромин.

Выдающиеся работы русских синтетиков, органиков XIX века, давшие много для развития органической химии, не находили применения у нас, а внедрялись в промышленность других стран, главным образом в Германии. Прекрасные работы Н. Н. Зинина по восстановлению нитросоединений, имеющие громадное значение для производства азокрасителей, были использованы впервые не в России, а в Германии.

Такой разрыв между научными достижениями и техникой мог оказаться опасным и даже губительным в годины бедствий. Война 1914—1917 гг. это ясно показала. Очень быстро были истрачены накопленные перед войной запасы, в армии и в тылу появилась острая нужда в необходимейших лекарственных препаратах. Нехватало взрывчатых веществ, порохов, снаряжения и вооружения. Для изменения этого положения уже в конце 1914 г. был основан Комитет по организации производства медикаментов в России, сделавший очень много для того, чтобы в стране появились такие необходимые соединения, как аспирин, салол и другие салициловые препараты, а также морфий, кодеин, стиптицин, кофеин, теобромин, новокаин и ряд других необходимых лекарственных средств.

С 1915 г. начинается освоение производства опийных алкалоидов, главным образом морфия и кодеина. Дружная совместная работа преподавательского и студенческого персонала Московского высшего технического училища быстро дала результаты. Было организовано сначала в полузаводском, а затем в заводском масштабе выделение морфия, кодеина и наркотина из опия, а также кофеина из чайной пыли и теобромина из шелухи какао. В это тяжелое время приходилось одновременно и работать в заводских условиях и проверять применяемые способы в лаборатории, а также разрабатывать методы аналитического контроля. Работы теоретического и прикладного характера шли одновременно и плодотворно. Создавались кадры молодых химиков, смело бравшихся и успешно осваивавших создание сложнейших фитохимических препаратов. В это время были разработаны превосходные методы получения синтетического кодеина, очистки и окисления наркотина, дан лучший метод

разделения важнейших опийных алкалоидов, усовершенствовано выделение теобромина и налажено сложное производство атропина из семян дурмана и из мандрагоры.

Впервые в нашей стране были открыты две специальные кафедры, имевшие большое значение для химико-фармацевтической промышленности: 1) кафедра химии и химической технологии фармацевтических препаратов в МВТУ и 2) кафедра химии алкалоидов во 2 МГУ (1920 г.).

Влияние работы МВТУ и 2 МГУ на теоретические и практические изыскания в фармацевтической химии было очень велико, оно ощущается и до сих пор.

Затем в развитии алкалоидной химии начинается некоторый перерыв. Во время гражданской войны и интервенции, когда юг и Туркестан были отрезаны, прекратился подвоз опия, а также мандрагоры (с Кавказа), из которой было начато успешное приготовление гиосциамина; не стало сырья для кофеина, для теобромина. Однако такое положение продолжалось недолго; в конце 20-х годов начался новый расцвет химии алкалоидов.

Изучение алкалоидов — этих интересных соединений переходит в Научно-исследовательский институт имени Орджоникидзе, где во главе алкалоидного отдела становится Александр Павлович Орехов, один из образованнейших русских химиков-органиков. Начав свое образование в России, А. П. Орехов не сразу определяет свое призвание. В 1899—1901 гг. он — студент Института путей сообщения в Петербурге; с 1901 по 1905 г. он учится в Горном институте в Екатеринославе, откуда исключается в связи со студенческими волнениями; с 1905 по 1908 г. он работает в Германии, сначала в университете в Иене у Кнорра, а затем в Гиссене у Наумана. В Гиссене Орехов защищает диссертацию на тему «Действие пятисернистого фосфора на ароматические амины» и получает звание доктора философии.

С 1910 по 1916 г. А. П. Орехов работает в Женеве ассистентом у проф. Гюи, затем переезжает в Париж, где ведет большую работу в лаборатории проф. Тиффно, продолжая главным образом начатые им в Женеве исследования по вопросу об интрамолекулярных перегруппировках. Этот период был очень продуктивным: им было опубликовано около 40 статей в журнале Французского химического общества и в Докладах Французской Академии наук.

В 1928 г. А. П. Орехов возвращается в Россию и начинает очень большую работу по изучению растительного сырья нашей страны, главным образом среднеазиатских республик.

Первый большой успех принесло исследование растения *Anabasis aphylla*, из которого им, совместно с Г. П. Меньшиковым, было выделено пять алкалоидов: анабазин, лупинин, афиллин, афиллидин и основание V. Из этих соединений важнейшим по свойствам оказался алкалоид анабазин $C_{10}H_{14}N_2$. Его строение, отвечающее α-пиперидил-β-пиридину, было безупречно установлено, а впоследствии подтверждено и синтезом, сделанным сотрудниками Орехова — Меньшиковым и Григоровским.

Анабазин — изомер никотина и, подобно последнему, обладает резко выраженными инсектисидными свойствами. Это его качество нашло практическое применение для борьбы с вредителями растений; до Великой Отечественной войны производство анабазина на Чимкентском заводе имени Дзержинского доходило до 140 т в год.

После выяснения строения анабазина и синтеза некоторых его производных, могущих, может быть, иметь и фармакологическое значение, А. П. Орехов перешел к изучению остальных алкалоидов *Anabasis*. Он показал, что между афиллидином и афиллином имеется генетическая связь, так как при каталитическом гидрировании афиллидин превращается в афиллин, присоединяя одну молекулу водорода. Оба алкалоида при электролитическом восстановлении переходят в одно и то же соединение, так называемый пахикарпин или d-спартеин.

Интерес к *Anabasis aphylla* сохранился и до сих пор. Сейчас это растение усердно изучается с химико-ботанической точки зрения учеником Орехова А. Садыковым в Ташкенте. Успешное исследование алкалоидов *Anabasis* имело большое значение для работы алкалоидного отдела. С этого времени начинается особенно деятельное изучение флоры наших среднеазиатских республик.

В сотрудничестве с Р. А. Коноваловой А. П. Орехов выделяет из растения *Convolvus pseudocantabricus* два новых алкалоида: конвольвин и конволамин, строение которых ими полностью устанавливается. Оба продукта оказались сложными эфирами вератровой кислоты и аминоспиртов нортропина и тропина. Так как оба аминоспирта входят в состав атропиновых алкалоидов, выделенных из семейства пасленовых, то вновь

было показано, что различные ботанические виды могут содержать близкие по строению алкалоиды:

атропин → сложный эфир тропина и троповой кислоты,

конволамин → сложный эфир тропина и вератровой кислоты,

конвольвин → сложный эфир тропина, нортропина и вератровой кислоты.

С фитохимической точки зрения было интересно также установление, что в семенах *Convolvus pseudocantabricus* главной составной частью является конвольвин, а в самом растении оба алкалоида находятся приблизительно в одинаковых количествах. Получается впечатление, что в процессе роста происходит переход конволамина в конвольвин, т. е. процесс деметилирования.

Оба алкалоида оказались местными анестезирующими средствами и, возможно, найдут и некоторое медицинское применение.

Оба аминоспирта нартропин и тропин сделались сравнительно доступными исходными веществами для различных синтезов, например для конвокаина и других сложных эфиров.

Очень большая работа была проведена А. П. Ореховым по различным видам *Sophora*. Из *S. pachycarpa* был выделен алкалоид пахикарпин $C_{15}H_{26}N_2$, оказавшийся по всем данным d-спартеином. Интересно, что семена этого растения не содержат пахикарпина, а в них находятся два других растительных основания: софокарпин и софокарпидин одинакового состава $C_{15}H_{24}ON_2$. Строение софокарпина не установлено, а софокарпидин после тщательного исследования оказался идентичным с матрином. Последний был выделен японским исследователем Кондо из *Sophora angustifolia*.

Из вида *S. lupinoides*, или, как его еще называют, *Thermopsis lanceolata*, было выделено пять алкалоидов: термопсин $C_{15}H_{20}ON_2$, гомотермопсин $C_{17}H_{24}ON_2$, анагирин $C_{15}H_{20}ON_2$, пахикарпин $C_{15}H_{26}N_2$ и метилцитозин $C_{12}H_{16}ON_2$. Особенно следует отметить отождествление алкалоида, первоначально обозначенного «алкалоид III», с анагирином и установление строения метилцитизина.

Из вида *S. alopecuroides* было найдено также пять алкалоидов: софорамин $C_{15}H_{20}ON_2$, софокарпин $C_{15}H_{24}ON_2$, матрин $C_{15}H_{24}ON_2$, алоперин $C_{15}H_{24}N_2$ и сафоридин $C_{15}H_{20}ON_2$. Их выделение и разделение было сопряжено с большими экспериментальными трудностями, тем не менее все они получены в аналитически чистом виде. Строение этих соединений не выяснилось.

Теоретический и практический интерес имеют работы А. П. Орехова и Н. Ф. Проскуриной по выделению из туркменистанских растений *Salsola Richteri* и *Salsola subaphylla* алкалоидов сальсолина и сальсолидина. Оба соединения оказались производными изохинолина; строение их было безупречно установлено и аналитически и синтетически. Синтез сальсолидина проведен А. П. Ореховым совместно с австрийским ученым Шпетом, одним из лучших знатоков химии алкалоидов; было показано, что это соединение является 1-метил-6-окси-7-метоксиизохинолином. Сальсолидин оказался *о*-метилированным производным сальсолина. Последний препарат благодаря своей способности устойчиво понижать кровяное давление нашел медицинское применение и в настоящее время изготовляется на Чимкентском химико-фармацевтическом заводе.

Небольшая, но изящная работа была сделана А. П. Ореховым по исследованию *Ammodendron Conollyi*, из которого был получен алкалоид, названный аммодендрином. Он оказался довольно близким по строению к анабазину, а именно моно-N-ацетильным производным тетрагидрированного анабазина. В строении этого соединения не установлены только положения двойной связи.

Интересные исследования проведены А. П. Ореховым при выделении алкалоидов из некоторых видов *Senecio*. Из этих работ надо остановиться на исследовании *S. platiphyllus*. Здесь были получены два новых алкалоида: сенецифиллин $C_{18}H_{23}O_5N$ и платифиллин $C_{18}H_{23}O_5N$. Строение этих соединений пока полностью не выяснено. Оба они оказались сложными эфирами; при гидролизе они дают аминоспирты. Сенецифиллин оказался эфиром известного уже аминоспирта ретроницина, получаемого при гидролизе очень многих других алкалоидов, выделяемых из различных видов *Senecio*. Платифиллин дал совершенно отличный насыщенный аминоспирт — платинецин. Алкалоид платифиллин приобретает некоторое фармакологическое значение, так как он может заменить атропин и гиосциамин при некоторых кишечных заболеваниях; он не вызывает неприятных побочных явлений, наблюдающихся при применении атропиновых препаратов. Незадолго до войны платифиллин был подробно исследован в фармакологическом отношении и был принят Медицинским советом. В настоящее время налаживается его производство.

Эти исследования в согласии с работами других авторов (Манске, Коновалова, Меньшиков, Баргер и др.) показали, что в основном алка-



лоиды видов *Senecio* при гидролизе дают аминоспирт ретронецин, тогда как алкалоиды видов гелиотропа — гелиотрин и лазиокарпин — при гидролизе расщепляются на гелиотридин и карбоновую кислоту. Однако материнским веществом в этих аминоспиртах является одно и то же гетероциклическое кольцо — пирролизидин. Отсюда следует, что эта группа растительных оснований образует одно естественное семейство, члены которого различаются между собой положением гидроксильных групп в азотосодержащем кольце и природой кислот, сложными эфирами которых они являются.

Незадолго перед преждевременной кончиной А. П. Ореховым было начато очень интересное и ценное исследование растений из семейства маковых, главным образом из видов *Papaver Armeniacum*, *P. floribundum* и *P. orientale*. Выделено несколько новых алкалоидов, из которых строение двух, ремерина и армепавина, было полностью установлено. Ремерин оказался метилендиоксиапорфином (I), а армепавин 4-оксибензильным производным N-метилдиметокситетрагидроизохинолина (II)

В поисках антималярийных средств А. П. Ореховым и его сотрудниками проведено много изысканий и синтетических работ по облагораживанию алкалоидов из доступного для нашей страны растения *Peganum harmala*. Однако эти исследования не дали положительных результатов. Вопреки довольно распространенному мнению, ни гармалин, ни гармин — главная составная часть экстракта, ни их производные, синтезированные и А. П. Ореховым с сотрудниками и другими исследователями, никакими антималярийными свойствами не обладают.

За московский период работы А. П. Ореховым сделано очень много. Трудно даже перечислить все исследования, проведенные им со своими сотрудниками. Помимо того, что было выделено много новых и интересных

соединений, нашедших практическое применение (анабазин, сальсолин, конвокаин, платифиллин и др.), им особенно ярко было подчеркнуто, что распространенное мнение о характерности нахождения в ботаническом виде только какого-нибудь определенного алкалоида или нескольких родственных между собой растительных оснований не может считаться безусловно верным. Не редко один и тот же алкалоид может встречаться в самых различных ботанических семействах. Так, в работе, проведенной им с С. Юнусовым и Коноваловой, было показано, что в растении *Glaucium fimbrilli gerum* не был найден характерный для других видов *Glaucium* глауцин, а совершенно неожиданно было установлено наличие хелеритрина — алкалоида, характерного для видов *Chelidonium Majus* (чистотел). В особенности интересно выделение Ореховым в сотрудничестве с теми же авторами эфедрина и псевдоэфедрина — алкалоидов, характерных для семейства хвойниковых, из *Roemeria refracta*, принадлежащего к семейству маковых.

А. П. Ореховым создана своя школа в Химико-фармацевтическом институте; из нее вышло очень много талантливых учеников, с успехом работающих в трудной области химии алкалоидов. Достаточно указать на таких превосходных работников, как постоянная сотрудница А. П. Орехова Р. А. Коновалова, давшая очень много ценного в области химии алкалоидов *Senecio* и различных видов *Papaver*. Превосходные исследования проведены Г. П. Меньшиковым с алкалоидами *Anabasis* и со сложными основаниями из некоторых видов *Heliotropium* и *Trachelanthus*. Из более молодых сотрудников можно назвать таких выдающихся исследователей, как Н. Ф. Проскурина, М. С. Рабинович, С. Юнусов, В. В. Киселев, А. Садыков.

Отдел, возглавлявшийся А. П. Ореховым, продолжает усиленно и успешно работать. Из отчета о проведенных Химико-фармацевтическим институтом за последние 20 лет работах видно, что отделом было обследовано более 900 растений нашей страны, среди них оказалось около 150 алкалоидоносов; выделено и описано 65 новых алкалоидов; выяснено строение около 20 алкалоидов; некоторые из них нашли медицинское применение.

Таким образом, начавшиеся под руководством А. П. Орехова работы внесли очень много нового и существенного в сокровищницу алкалоидной химии. Вклад этот очень значителен, если принять во внимание,

что за время от 1817 г. и до наших дней описано всего около 600 алкалоидов.

Во время Отечественной войны 1941—1945 гг. сильно затормозилось развитие дисциплин, не имевших непосредственного значения для военных нужд. Все силы были направлены на борьбу с жестоким врагом. Начавшийся мирный организационный период вновь создал благоприятные условия для научно-исследовательских работ. Налаживаются ботанические экспедиции по необъятной советской земле, организуется сбор лекарственного растительного сырья, улучшается положение с органическими растворителями, нужными для извлечения алкалоидов.

Можно быть уверенным, что исследования в области химии алкалоидов, так успешно начатые А. П. Ореховым, будут не менее успешно продолжаться его сотрудниками и учениками и найдут отклик и в других институтах и лабораториях нашей родины.

Академик
В. М. Родионов

ИССЛЕДОВАНИЯ В ОБЛАСТИ БЕТА-АМИНОКИСЛОТ

За истекшие сто лет органическая химия сделала громадные успехи как в изучении теоретических основ науки, так и в развитии промышленного органического синтеза. Замечательные синтетические работы Велера, Бертло, Кольбе, Зинина, Бутлерова, Кекуле, Дюма, Перкина, Адольфа Байера, Гребе, Либермана, Грисса, Костанецкого, Э. Фишера, а в более позднее время — Кижнера, Демьянова, Зелинского, Фаворского, Рабе, В. Перкина, Р. Робинсона, Питке, Вильштеттера, Каррера, Удуорда, Деринга (синтез хинина), Дю-Виньо и многих других выдающихся исследователей показали, что самые сложные природные органические вещества могут быть приготовлены лабораторным, а затем и фабричным путем.

Особенно большое значение в развитии органического синтеза имело создание Бутлеровым и Кекуле общепринятой ныне теории строения органических соединений. Самое важное в этой теории это то, что она дает возможность предугадывать беспредельность научных открытий. Основываясь на этой теории, работники органической химии шли от успеха к успеху. Синтезы таких технически ценных веществ, как ализарин, индиго, органические красители и душистые вещества, нашли применение в промышленности и очень быстро вытеснили более дорогие и зачастую более трудно поддающиеся очистке природные соединения.

Особенно усиленно промышленный органический синтез начал развиваться в XX веке. Появились новые мощные отрасли промышленности, имеющие большое народнохозяйственное значение. Синтез высокомолекулярных соединений повел к созданию производства пластических масс, во многих случаях с успехом заменяющих дорогие природные смолы, изоляторы, бальзамы, металлы, а в последнее время даже природные текстильные материалы. Углубленное изучение нефти повело к выявлению новых источников промышленного сырья для приготовления высокооктанового моторного топлива, производства синтетического каучука, получения ароматических углеводородов, а также для многих продуктов тяжелого органического синтеза.

За последнее время отмечен все более усиливающийся уклон органической химии в биологию, многие насущные вопросы которой могли быть разрешены лишь при помощи органической химии. Установлено строение многих физиологически важных соединений, например, хлорофилла, красящих веществ крови и желчи, гормонов (в том числе и гормонов пола), желчных кислот, стеринов, витаминов. Совсем недавно было начато совместное изучение таких биологически и химически важных соединений, как антибиотики (вещества типа пенициллина, стрептомицина, грамицидина и др.).

Многие из недавно открытых соединений уже синтезированы и вытеснили природные продукты. Витамины B, C, D, никотиновая кислота и некоторые другие готовятся в фабричном масштабе и обходятся дешевле естественных соединений.

Из трех основных классов веществ, входящих в состав пищи млекопитающих и человека, два класса — углеводы и жиры — глубоко изучены и могут получаться фитохимическим и синтетическим путем. Некоторые из них уже готовятся на фабриках. Например, глюкоза получается гидролизом из крахмала и дешевой древесины и служит кормовым средством для скота и исходным материалом для получения винного спирта.

Совсем недавно германская промышленность разработала методы приготовления пищевых жиров из продуктов окисления синтетически получаемых парафинов и рекомендует эти жиры диабетикам, как совершенно не содержащие углеводов. Замена природных жиров искусственными препаратами — вопрос лишь калькуляции.

Меньшие успехи могут быть отмечены при изучении белковых веществ протеинов, третьего важнейшего класса пищи и вообще существенной части всякой живой материи. Вместе с водой, солями и липоидами белковые вещества образуют материал, из которого построена протоплазма животных и растительных клеток. Обычные пути выяснения химической природы белков оказались недоступными из-за неустойчивости белкового материала и невозможности получения его в чистом виде.

Белки не перегоняются без разложения, не плавятся и за очень редкими исключениями не кристаллизуются. К концу XIX века очень многие замечательные исследователи считали изучение белков совершенно неблагодарной задачей. Один из самых выдающихся органиков того времени, автор многих сложнейших синтезов, Адольф Байер высказал в 1895 г. такую мысль: «До сих пор нет никакой надежды в ближайшем будущем определить природу белка. Должны ли мы задаваться подобными вопросами, если возможность их решения совершенно невероятна, или уже будем продолжать работу в тех областях, где мы наверное соберем богатую или хотя бы скромную жатву. Я предпочитаю последнее». Однако не прошло и пяти лет, как началось усиленное изучение белковых веществ, и в настоящее время огромное число исследований посвящено как раз белковой проблеме.

К разрешению вопроса о строении протеинов, по инициативе Э. Фишера, одного из самых ближайших учеников Байера, подошли с другой стороны. Сложная молекула (или молекулярный комплекс) белковых веществ была подвергнута Фишером кислотному гидролизу и было установлено, что эти соединения — продукт конденсации очень большого числа различных α-аминокислот, т. е. соединений типа $RCH(NH_2)COOH$. Первоначально попытки найти другие аминокислоты ($\beta \rightarrow RCH(NH_2)CH_2COOH$, $\gamma \rightarrow RCH(NH_2)CH_2COOH$ и др.) в протеинах не увенчались успехом, а потому все внимание исследователей было сосредоточено исключительно на изучении α-аминокислот. Интерес к ним особенно возрос, когда Э. Фишеру удалось показать, что взаимной конденсацией этих соединений можно готовить белковоподобные вещества, так называемые полипептиды, во многих физических и химических отношениях сходные с природными белками. Была тщательно исследована химическая природа α-аминокислот, разработаны доступные методы их синтетического получения; с ними было выполнено очень много разнообразных превра-

щений, ценных как с химической, так и биологической точки зрения. Бесспорно было установлено, что α-аминокислоты, возникающие в живом организме в результате энзиматического распада, в свою очередь служат в нем строительным материалом для синтеза важнейших биологических веществ, главным образом протеинов.

Изучение остальных классов аминокислот (β, γ и др.) проводилось в очень небольших масштабах, ибо полагали, что эти соединения, не входящие в белковый комплекс, могут иметь лишь второстепенное научное и биологическое значение. Не было и хороших способов их получения.

Постепенно, однако, положение с вопросами о β-аминокислотах начало резко меняться. Сейчас уже нет уверенности, что β-аминокислоты не входят в состав белкового комплекса. Возможно, что самая методика выделения аминокислот из протеинов, сводящаяся к кислотному гидролизу с последующим превращением смеси аминокислот в хлоргидраты их эфиров и дробной разгонкой их при очень большом разрежении, есть причина потери β-аминокислот. Эти соединения мало стойки к действию крепких минеральных кислот, легко теряют аминогруппу, переходят в ненасыщенные соединения и потому не могут быть выделены при действии хлористого водорода и теряются в маточных растворах.

Однако тщательно проведенным исследованием В. С. Гулевича было показано, что карнозин (I) и ансерин, дипептиды простейшей β-аминокислоты — β-аланина и гистидина (или соответственно N-метилгистидина), встречаются в мышечных тканях различных животных (млекопитающие, рыбы, пресмыкающиеся, птицы)

$$\underline{H_2N-CH_2-CH_2-CO.}NH-\underset{\displaystyle COOH}{\underset{|}{CH}}-CH_2-\underset{\displaystyle NH-CH}{\underset{|}{C}}=CH\!\!\!\!\!\!\!\!\!\;\;\;\;\;\;\;\;\; \!\!>\!N \qquad (I)$$

(подчеркнут остаток β-аминокислоты).

Таким образом, впервые было установлено, что в состав экстрактов белкового комплекса могут входить и производные β-аминокислот.

Еще большее впечатление произвело выделение из самых разнообразных животных тканей так называемой пантотеновой кислоты, продукта конденсации опять-таки β-аланина и диметилдиоксимасляной кислоты (II)

$$HO-H_2\overset{\displaystyle CH_3}{\overset{|}{\underset{\displaystyle CH_3}{\underset{|}{C}}}}-CH(OH)-CO.\underline{NH\,CH_2CH_2COOH} \qquad (II)$$

Это соединение оказалось универсальным витамином, необходимым для роста всех живых существ: бактерий, водорослей, семян растений, всех низших и высших животных, и, повидимому, человека. В 1940 г. это соединение было синтезировано из диоксидиметилмасляной кислоты и β-аланина. Изучение свойств пантотеновой кислоты показало, что существенная часть этого соединения состоит как раз из β-аминокислоты. При попытках заменить β-аланин α-аминосоединениями получается вещество, не имеющее витаминного действия. Естественно было испытать действие самого β-аланина в этом отношении. Многочисленными работами было установлено, что в некоторых случаях сам β-аланин имеет большое значение для развития микроорганизмов. В особенности резко сказывается его стимулирующее влияние на рост различных видов дрожжей. Возможно, что действие β-аланина сводится к тому, что в микроорганизмах β-аланин превращается в пантотеновую кислоту.

При сравнении биологического поведения β-аминокислот с α-аминокислотами найдено, что они во многом различаются между собой. Судьба β-аминокислот в организме, повидимому, различная. Они легко изменяются, возможно дезаминируются, переходят в кетосоединения и в этом виде, может быть, даже вредны для организма (ацетонурия); они могут также легко переходить (как это установлено in vitro в наших работах) в сложные пиримидиновые, а может быть, и в пуриновые соединения и в таком виде входят в состав сложнейших белковых веществ — нуклеопротеидов.

Понятен поэтому все возрастающий интерес к этим веществам; число работ по β-аминокислотам растет с каждым годом.

Изучение этих соединений очень облегчилось в связи с тем, что в 1926 г. Родионову и его сотрудникам Малевинской и Федоровой удалось найти очень простой и доступный способ получения этих соединений путем конденсации ароматических альдегидов с малоновой кислотой в присутствии спиртового раствора аммиака

$$RCHO + NH_3 + CH_2(COOH)_2 \rightarrow RCH(NH_2)CH_2COOH + CO_2 + H_2O .$$

Попутно получающееся производное коричной кислоты

$$RCHO + CH_2(COOH)_2 \rightarrow RCH = CHCOOH + CO_2 + H_2O$$

легко может быть отделено, и β-аминокислоты по этому способу легко получаются в очень чистом виде.

В дальнейшем Родионовым в сотрудничестве со Зворыкиной и Ярцевой было показано, что этот способ может быть с успехом распространен и на алифатические альдегиды. Таким образом, была установлена общность найденной реакции. Также не безуспешно эта реакция была распространена на некоторые гетероциклические альдегиды, что снова подтвердило универсальность найденного метода. β-аминокислоты сделались легко доступными препаратами, что повело к постановке многих исследований по превращению этих соединений в некоторые биологически интересные вещества, главным образом в производные пиримидина и имидазола.

Такие работы производились не только у нас, но и за рубежом. Известный американский исследователь Т. Джонсон для синтеза пиримидиновых соединений предпочел взять наш способ синтеза β-аминокислот — исходного материала для приготовления пиримидинов.

Гексагидрированные пиримидины (или дигидроурацилы) получаются из β-аминокислот с очень хорошими выходами при их взаимодействии с алкалицианатами или при нагревании с мочевиной

$$\underset{\displaystyle NH_2}{RCH}-CH_2COOH + NH_2CONH_2 \rightarrow RCH\begin{matrix} \diagup NH \diagdown \\ \\ \diagdown CH_2 - CO \diagup \end{matrix}\!\!\begin{matrix} CO \\ | \\ NH \end{matrix} + NH_3 + H_2O$$

Эти соединения по строению очень близки к так называемым урацилам, тетрагидрированным пиримидинам, продуктам расщепления биологически важных нуклеиновых кислот. В развитии этих работ удалось приготовить и тетрагидропиримидины, т. е. соединения еще более близкие к урацилам.

В. К. Зворыкиной было показано, что амиды ацилированных β-аминокислот с легкостью превращаются в тетрагидропиримидины при их взаимодействии с хлористым тионилом с последующей обработкой получаемых хлорангидридов сухим аммиаком. Повидимому, реакция протекает по такой схеме:

$$RCH\begin{matrix} \diagup NH-COR' \\ \diagdown CH_2-COOH \end{matrix} + 2SOCl_2 \rightarrow RCH\begin{matrix} \diagup N{=}C\begin{matrix} \diagup R' \\ \diagdown Cl \end{matrix} \\ \diagdown CH_2-COCl \end{matrix} \xrightarrow{+NH_3} RCH\begin{matrix} \diagup N{=}C(R') \diagdown \\ \qquad\quad NH \\ \diagdown CH_2 - CO \diagup \end{matrix}$$

И этот способ оказался общего характера и дал возможность готовить всевозможные тетрагидропиримидиновые соединения, несомненно имеющие большое биологическое, а возможно, и фармакологическое значение.

О важности этих работ можно судить по тому, что такие ценные общеизвестные препараты, как веронал и лучший сульфамид, сульфодиазин, относятся к классу пиримидиновых соединений. В последнее время намечается и возможность превращения β-аминокислот в дигидропиримидины и просто пиримидины.

Выше было указано, что простейшая β-аминокислота — β-аланин — сильно увеличивает рост низших организмов (различных видов дрожжей). Отсюда — повышенный интерес к этому соединению и многочисленные попытки новых синтезов этой β-аминокислоты во многих странах (США, Канада, Швейцария, Германия). В 1946 г. в нашей лаборатории удалось разработать очень простой способ получения этого соединения из вполне доступного сырья (Н. Г. Ярцева).

Эфиры акриловой кислоты в присутствии гидроокиси триметилфениламмония в качестве катализатора с превосходным выходом (около 95—96%) присоединяют фталимид с образованием эфиров фталимидопропионовой кислоты. Последние при нагревании с минеральными кислотами дают почти количественно соли β-аланина

$$C_6H_4\left\langle\begin{matrix}CO\\CO\end{matrix}\right\rangle NH + CH_2 = CH.CO_2R \rightarrow C_6H_5\left\langle\begin{matrix}CO\\CO\end{matrix}\right\rangle N - CH_2 - CH_2 - CO_2R,$$

$$C_6H_4\left\langle\begin{matrix}CO\\CO\end{matrix}\right\rangle N.CH_2 - CH_2CO_2R + HCl + 3H_2O \rightarrow$$

$$\rightarrow HClNH_2CH_2CH_2COOH + ROH + C_6H_4(COOH)_2.$$

Определенный интерес представило испытание и других доступных β-аминокислот на влияние их на рост низших организмов. Опыты показали, что некоторые β-аминокислоты, например синтезированная В. К. Зворыкиной β-аминопеларгоновая кислота, обладает стимулирующим действием на рост дрожжей, тогда как другие, например β-аминомасляная кислота, скорее даже угнетают развитие микроорганизмов. Это ведет к постановке совместных с микробиологами исследований по выяснению тонких различий в строении аминокислот, обусловливающих принципиально важное отличие в их поведении по отношению к микроорганизмам.

При расширении этих работ было установлено, что одна из простейших ароматических аминокислот, так называемая антраниловая кислота, тоже обладает, хотя и слабым, стимулирующим действием на рост дрожжей. Это в свою очередь очень расширяет круг исследований и ведет к постановке поисковых работ во многих других областях.

Тщательные работы по изучению гофмановской реакции с амидами β-аминокислот дали два очень интересных результата. Первый состоит в том, что вновь бесспорным путем было установлено, что образование аминов из амидов карбоновых кислот идет через изоциановые эфиры

$$RCONH_2 \rightarrow RNCO \xrightarrow{H_2O} RNH_2 + CO_2 .$$

Вторым результатом является синтез биологически активного вещества с биотинным действием.

Приготовляемый из амида β-аминопеларгоновой кислоты изоциановый эфир (I) претерпевает внутримолекулярную перегруппировку с образованием гексилглиоксилидона (II), кольца, входящего в молекулу биотина (III) и дестиобиотина (IV)

```
              NHCOR      NH — CO — NH     NH — CO — NH      NH — CO — NH
C6H13 CH<            →   |         |      |         |       |          |
              CH2NCO     CH ——————— CH2,  CH ——————— CH     CH ———————— CH
    (I)                  |                |         |       |           |
                         (CH2)5           CH        CH2     (CH2)5      CH3
                         |                | \      /        |
                         CH3              |   \  /          COOH
                                          |    S
                         (II)             (CH2)4                (IV)
                                          |
                                          COOH
                                          (III)
```

Близкое сходство по строению соединений (II) и (IV) навело на мысль исследовать их микробиологически. Было найдено (М. Н. Майсел, Микробиологический институт АН СССР), что гексилглиоксилидон (II) очень близок по свойствам к дестиобиотину (IV). Это установление имеет очень большое значение и ведет, с одной стороны, к постановке работ по изучению связи между строением произведенных превращений в молекулах β-аминокислот и их биотинным действием, а с другой — открывает возможность подойти к синтезу дестилбиотина и, может быть, самого биотина, т. е. одного из самых существенных биологических факторов.

В настоящее время многочисленными работами установлено, что биотин (витамин Н) — необходимая составная часть диэты для всех видов животных (по крайней мере для всех до сих пор изученных видов). В области микроорганизмов биотин — существенный фактор роста для различных рас дрожжей, бактерий, водорослей и т. п.

Таким образом, благодаря простому и изящному способу получения β-аминокислот, разработанному в нашей лаборатории, получилась возможность синтезов таких ценных и интересных соединений, как пиримидины и их ди-, тетра- и гексагидропроизводные, а также приготовления веществ, обладающих биотинным действием, и многих других биологически ценных препаратов.

Член-корреспондент АН СССР
А. П. Виноградов

В. И. ВЕРНАДСКИЙ И ГЕОХИМИЯ РЕДКИХ ЭЛЕМЕНТОВ

Владимир Иванович Вернадский — это целая эпоха в развитии нашей науки. В течение своего полувекового упорного творческого труда он открыл в науке новые области знания и оказал глубочайшее влияние на весь ее ход.

Это был крупнейший естествоиспытатель и мыслитель, поражавший каждого, кто соприкасался с ним, исключительной широтой своих интересов, богатством и глубиной знания в самых разнообразных областях: минералогии и кристаллографии, геохимии и биогеохимии, геологии и радиогеологии, химии, биохимии и почвоведения, гидрологии, истории знания и философии. Он оставил нам огромное научное наследство в обширном кругу своих учеников и сотрудников, в созданных им в нашей стране многочисленных научных институтах, в написанных им прекрасных книгах. В. И. Вернадский бесспорно является основоположником новой науки — геохимии.

В 1908 г. выходит в свет известная монография Ф. У. Кларка «Data of Geochemistry», где автор впервые делает попытку обобщить все известные числовые данные по химическому составу земной коры, пород, океана и т. д. и где впервые в нашем столетии появляется самое слово «геохимия». Однако Кларк, собрав многочисленные данные, не сделал соответствующих выводов, не поставил перед геохимией каких-либо задач. Это было вскоре сделано В. И. Вернадским, с предельной ясностью сформулировавшим основные задачи новой науки, а именно: изучение

истории земных атомов, исследование поведения химических элементов в земной коре.

Мы должны, конечно, заметить, что это стало возможным лишь благодаря новой атомистике, которая уже успела к этому времени проникнуть в разнообразные области химического и физического знания. Вместе с тем этот переход на рельсы новой науки был подготовлен проделанной В. И. Вернадским в стенах Московского университета (1890—1911) огромной работой по созданию генетической, динамической минералогии. Идеи геохимии были очень живо восприняты. У В. И. Вернадского они воплощены были позже в блестящих работах по строению силикатов и алюмосиликатов земной коры (1891—1937), в многотомных, непревзойденных трудах «Опыт описательной минералогии» (1908—1922) и «История минералов земной коры» (1923—1934), в замечательных «Очерках геохимии» (1924—1929—1934), в многочисленных исследованиях по геохимии биосферы, природных вод, газов, живого вещества, по радиоактивности, метеоритике и в других трудах, хорошо знакомых и неспециалистам.

В настоящей статье было бы невозможно полно и критически рассмотреть всю сумму знания, которую В. И. Вернадский внес в созданную им геохимию. Поэтому мы несколько подробнее остановимся лишь на одной из многочисленных геохимических задач, которая была выдвинута талантом В. И. Вернадского и блестяще оправдана всей практикой нашей жизни. Мы имеем в виду его исследования по геохимии так называемых «редких» химических элементов. (Сам Владимир Иванович избегал этого названия.) Это интересно и важно, как нам кажется, потому, что он, как и в ряде других аналогичных случаев, предвидел значительно раньше других практическое значение этих химических элементов, и еще потому, что он своими исследованиями в этом направлении заложил прочное основание работам, размах которых мы сейчас у себя наблюдаем.

В 1902 г. В. И. Вернадский вместе с С. П. Поповым впервые открывает буру в России — в грязи и водах керченских грязевых вулканов. Он правильно заключает, что бор находится, повидимому, во всех грязевых вулканах. Это вызвало в свою очередь большую серию исследований, проделанных многими нашими учеными в связи с поисками бора в России, с попыткой использовать с этой целью подземные минеральные воды. Эти работы не замирают даже и сейчас после открытия Индерского боратного месторождения (1932).

Работы по бору были небольшим эпизодом в исследованиях В. И. Вернадского и стоят несколько в стороне от серии других задуманных им в это время работ по геохимии редких элементов. С 1908 г. в Минералогическом музее Академии Наук в Петербурге он ставит задачу по систематическому изучению распространения в земной коре рубидия, цезия, лития, индия, таллия и других редких элементов. Он ищет метод, который дал бы ему возможность быстро накопить значительное число данных. Бывая за границей, он видел спектральные исследования минерального сырья, поставленные у де-Грамона, Юрбена. Но в России никто до него таких исследований не вел. Он останавливается на спектральном методе, который в то время далеко еще не был разработан и давал лишь полуколичественную оценку. В. И. Вернадский со своими учениками Е. Д. Ревуцкой, Б. А. Линденером и А. А. Твалчрелидзе организует в Минералогическом музее АН спектральные работы, комбинируя их с химическими анализами. Он приобретает прибор Хильгера, а несколько позже заказывает специальный прибор с диффракционной решеткой, имеющий метровую шкалу. Эти работы систематически велись им с 1908 по 1916 г. Так впервые у нас В. И. Вернадский ввел спектральный анализ в геохимическую практику и предвосхитил дальнейшее его блестящее развитие в этом же направлении. Но вернемся к геохимическим результатам этих исследований.

Первая попытка В. И. Вернадского дать геохимическую картину поведения химического элемента в земной коре относится к скандию. Эбергард в 1908 г. опубликовал многочисленные определения скандия в минералах и породах, но в заключение написал, что он не мог заметить какойлибо закономерности в распределении скандия. В. И. Вернадский пересматривает весь известный ему материал и показывает, что скандий сосредоточивается в участках земной коры, связанных с выходами кислых магм, выделяющихся при застывании в пневматолитической стадии, т. е. в пегматитах. Действительно, вскоре в гранитных пегматитах был открыт настоящий минерал скандия. В 1931 г. Гольдшмидт, развивая эти работы, обнаруживает, что значительное количество скандия остается в магниевых и железных силикатах ранней и главной кристаллизации магмы. Так начал свои работы В. И. Вернадский по геохимии.

Но он быстро расширил круг изучаемых им элементов. Определения рубидия, цезия, лития, таллия, индия, отчасти галлия и др. в минералах

и породах спектроскопически можно было вести одновременно. Собирается большой материал. Он позволяет надежно подытожить наблюдения. В. И. Вернадский выступает с геохимическими очерками по распределению рубидия, индия и таллия. К геохимии рубидия он возвращался несколько раз (1909, 1911). Он обнаруживает высокое содержание рубидия (и цезия) в полевых шпатах, микроклине, лепидолитах и других слюдах и т. д., как правило — в калиевых алюмосиликатах. Между тем Барбье, который вел исследования по рубидию, выступил с утверждением, что рубидия в микроклинах нет и что он встречается только в ортоклазах. В. И. Вернадский химическим путем выделил из микроклина с Ильменских гор 3.12% Rb_2O. Эти прекрасные исследования были позже у нас же подтверждены многократно (Мальчевская-Рутковская, Толмачев и Филиппов и др.). В. И. Вернадский показал, что рубидий (и цезий) в своей геохимической истории следует за калием. Он впервые делает попытку установить отношение между калием и рубидием в породах и минералах, вводя тем самым очень важный в современной геохимии принцип пропорциональности или постоянства отношения между парой близких по свойствам химических элементов. Ученики и сотрудники В. И. Вернадского пополнили наши знания по геохимии рубидия (Е. С. Бурксер, 1936; Т. Ф. Боровик-Романова, 1946 и др.). Уже в 1916 г. делается попытка получения рубидия, цезия и лития из богатого сырья лепидолитов Липовки на Урале, из Забайкалья, для появившихся тогда фотоэлементов. В значительном масштабе это осуществляется, однако, только в 1931 г. (Бурксер и др.). Геохимия лития не могла быть хорошо изучена, так как литий встречался в рассеянном состоянии во всех исследованных материалах. Было трудно проследить пути его концентрации. Позже П. П. Пилипенко, также ученик В. И. Вернадского, дал прекрасный очерк о литии и обнаружил его новые накопления в главконите, в боратах и т. д.

На основе собственных данных и материалов других авторов В. И. Вернадский пишет далее геохимический очерк по индию. Эта работа полностью сохранила свое значение до наших дней. Он указывает, в частности, что индий относится к изоморфному ряду цинк — кадмий — марганец — железо. Именно в цинковых обманках, в их первых генерациях, так называемых «мармотитах», наиболее древних, бедных галлием, содержащих железо и марганец, индий встречается в наибольших количествах. Он обращает внимание на парагенез индия с оловом, вольфрамом и т. д. Все эти на-



В. И. Вернадский

блюдения полностью оправдываются в работах сотрудников В. И. Вернадского (Н. Н. Прокопенко, С. А. Боровик, Н. И. Влодавец). В результате в 1931 г. индий был получен у нас из сульфидов цинка и из каттрельной пыли от обжига сульфидных руд.

После работ В. И. Вернадского по таллию, в которых он показывает изоморфное замещение калия (и, следовательно, рубидия) в минералах на таллий — в полевых шпатах и т. д., почти ничего нового по геохимии таллия не появлялось. Все эти исследования по распределению рубидия, цезия, лития, таллия и др. в земной коре шли у Владимира Ивановича под знаком двух идей — выяснения количественного состава земли и изучения закономерностей парагенезиса этих элементов. Еще в 1909 г. в блестящем докладе на съезде естествоиспытателей о парагенезисе химических элементов В. И. Вернадский говорит о первых своих обобщениях в этом направлении. В своих 18 природных изоморфных рядах он дает общую схему распределения химических элементов в земной коре. Изоморфные ряды В. И. Вернадского сыграли огромную роль в понимании законов распределения парагенезиса химических элементов. Он делает при этом чрезвычайно важное замечание, что эти изоморфные ряды, в зависимости от условий, изменяются. Число входящих в них химических элементов увеличивается при повышении температуры и давления.

Важно отметить и другое обстоятельство. Если изоморфные ряды позволяют обнаружить известные закономерности в распределении многих химических элементов, то в отношении некоторых элементов, например иода, галлия, германия и др., это удается лишь с трудом. Эти элементы присутствуют всюду в очень малых количествах. Они находятся в рассеянии. Так впервые формулируется принцип «всюдности» — в любом куске вещества, чистейшем кварце или каррарском мраморе и т. д., можно найти все химические элементы. Этот вопрос спустя 30 лет неожиданно вновь привлек внимание геохимиков.

В. И. Вернадский в своей геохимической классификации химических элементов выделяет группу рассеянных химических элементов, которые совсем уж не так редки, как обычно до него считали. Но они рассеяны и не образуют собственных минералов. Однако их масса, подчеркивает он, огромна. Вопрос о пределе рассеяния химических элементов в земной коре часто занимал Владимира Ивановича. В последний раз он вернулся к нему в речи к Общему Собранию Академии Наук в 1927 г. Он ясно видел, что

с этим рассеянием природное вещество приобретает новое свойство по сравнению с «чистым»: например, у сульфида кальция при содержании в нем следов некоторых рассеянных элементов вызывается люминесценция и т. п. явления. С другой стороны, рассеяние элементов возникает также и вследствие радиоактивного распада, в частности под влиянием космического излучения. В «Материалах по спектроскопии земной коры» (1916) он рассматривает концентрацию и рассеяние химических элементов как две стороны одного и того же процесса. В той же работе он ставит новые задачи по геохимии отдельных элементов — их количественный учет, изучение распределения в различных частях земной коры (региональная геохимия) и нахождение их в своеобразных химических областях (геохимических провинциях — как говорим мы теперь) как географически, так и батиграфически (геологически). К рассмотренным нами выше работам относится и последний, седьмой очерк этой серии работ по распространению висмута (1916). Страна переживает в это время войну, ощущает недостаток сырья. И как бы отвечая на это, В. И. Вернадский заканчивает свою геохимию висмута словами, что «для висмута нет области большего обогащения, чем область руд олова, золота, серебра, свинца, никеля, кобальта, и является легкомысленной непредусмотрительностью не извлекать попутно висмут, безвозвратно теряющийся при добыче этих минералов». Только в 1922 г. ученик В. И. Вернадского К. А. Ненадкевич получил из этого сырья висмут.

Обладая собственным опытом по распределению химических элементов в земной коре, имея перед собою свои «Опыты описательной минералогии», В. И. Вернадский, в связи с усиливающимся «минеральным голодом» в стране, выступает с критикой использования химических элементов в тогдашней России. Он показывает, что только 30 химических элементов используются у нас и лишь 17 — в достаточном количестве, т. е. наполовину меньше, чем за рубежом. Он неоднократно утверждал, что в России имеются все геологические формации и что в недрах территории России протекали и протекают все геохимические и рудообразовательные процессы, которые имеются в других частях земного шара. Он пишет в этой заметке: «Мы нигде не встречаемся здесь с химическими элементами, нахождение которых в удобной для разработки форме не было бы вероятным, т. е. не заслуживало бы поисков с надеждой на успех. Пять элементов, по отношению к которым, на первый взгляд,

дело обстоит наименее благоприятно, — гелий, неон, калий, кобальт и никель; но и по отношению к ним всякий знакомый с геологией или минералогией России найдет целый ряд ясных указаний на то, что дело идет не об отсутствии этих элементов, а о нашем недостаточном знании о них». Все это полностью оправдалось: в советский период были открыты Соликамск (1925), руды никеля и кобальта и многие другие.

Владимир Иванович организует в Академии Наук Комиссию по изучению естественных сил (КЕПС — 1915—1930 гг., ныне СОПС). Он возбуждает вопрос об исчерпывающем изучении всех спутников платины, требует прекращения вывоза сырой платиновой руды за границу, ставит вопрос о поисках сырья на алюминий, ванадий, селен, теллур и др. Он привлекает к составлению монографических очерков по распространению редких и других элементов огромный коллектив своих учеников и сотрудников. За короткий срок (1916—1930) выходят в свет первые монографические очерки по литию, барию, вольфраму, олову, бору, титану, ванадию, урану и т. д. В создании их принимают участие Я. В. Самойлов, В. Г. Хлопин, В. В. Аршинов, А. Е. Ферсман, Д. И. Щербаков и др. Эти работы, несомненно, сыграли исключительно важную роль в расширении возможностей по использованию в нашей стране редких и других химических элементов.

Владимир Иванович продолжает детально изучать геохимическую историю марганца, излагая ее в ряде классических работ (1924, 1934, 1937), детально разрабатывает геохимию титана (1910, 1935), редких земель (1929).

Вместе с тем молодая Советская Республика требует молибдена, хрома, никеля, кобальта для стали. Привлекаются десятки, сотни учеников В. И. Вернадского — минералогов, геохимиков — на поиски и разведку нужного сырья. Он, как всегда, спешит им помочь.

А. Е. Ферсман со своими учениками берется за разрешение новых геохимических задач и достигает больших успехов. С открытием кольских богатств под руководством А. Е. Ферсмана производятся замечательные исследования над распространением и парагенезисом совершенно новых для нас химических элементов: циркония (Костылева), редких земель (Старынкевич), ниобия (Герасимовский), а также по геохимии титана (Малышев), ртути (Сауков), бериллия (Власов).

Примерно с 1931 г., как бы соревнуясь с нами, на Западе Гольдшмидт начинает свои геохимические работы по редким элементам, пользуясь новейшими спектроскопическим и рентгено-спектральным методами.

За последние 10—15 лет в геохимии накапливаются значительные материалы по распространению химических элементов на земле. За рубежом и у нас накануне второй мировой войны в промышленность, в той или иной степени, вовлекаются все химические элементы (за небольшими исключениями).

В течение 1924—1934 гг. выходит ряд изданий «Очерков геохимии», где В. И. Вернадский пересматривает свою таблицу распространения химических элементов в земной коре. Его исчисления по декадам (число lg) глубоко входят в геохимическую практику. Он дает классификацию геохимических элементов. Его возраст приближается к 80 годам. Но он не только не отходит от этой проблемы, а еще более углубляется в изучение законов распределения в земной коре элементов последней строки менделеевской таблицы. Он часто вспоминал о том большом впечатлении, которое на него произвела речь Джоли о радиоактивной теории магматических процессов, произнесенная в Дублине в 1908 г. С этого момента во всех работах В. И. Вернадского заметно его настороженное внимание к явлениям радиоактивности. Еще в 1910 г. он проникновенно говорил: «В вопросе о радии ни одно государство и общество не могут относиться безразлично — как, каким путем, кем и когда будут использованы и изучены находящиеся в его владениях источники лучистой энергии. Ибо владение большими запасами радия дает владельцам его силу и власть, перед которой может побледнеть то могущество, какое получают владельцы золота, земли, капитала».

В 1914 г. ему, наконец, удается организовать экспедиции по исследованию радиоактивных месторождений России. В них участвуют лучшие силы — ученики Владимира Ивановича и его сотрудники — Пилипенко, А. Е. Ферсман, К. К. Матвеев, В. И. Крыжановский, К. А. Ненадкевич, В. С. Наливкин и многие другие. В. И. Вернадский сам совершает большой маршрут через Кавказ к предгорьям Алая, в Ферганскую долину, Забайкалье, на Урал. Это была первая глубокая научная разведка на радиоактивность. Ее результаты — основа всех дальнейших наших поисков радиоактивных элементов. К отчетам В. И. Вернадского, к его лучшей у нас сводке по геохимии радиоактивных элементов в «Очерках геохимии» мы постоянно возвращаемся. Он добивается, наконец, создания Радиевого института, привлекает к участию в его работах В. Г. Хлопина, собирает вокруг себя прекрасный, мощный коллектив. По его замыслу в Инсти-

туте организуется геохимический отдел, откуда выходят работы по геохимии радиоэлементов (Л. В. Комлев, И. Е. Старик), а также по геохимии благородных газов (В. Г. Хлопин). В. И. Вернадский обращает внимание на высокое содержание радия в нефтяных водах как на источник промышленного радия. Он увлекается происхождением тория и мезотория, открытого В. И. Барановым в природных водах. Шаг за шагом он стремится познать сложные пути миграции элементов последней строки менделеевской таблицы. Задачи, которые он себе ставил, поражают глубиной, а путь, которым он шел, — своей целеустремленностью. Для него не было сомнения в правильности намеченной цели.

В 1922 г. он пишет: «Мы подходим к великому перевороту в жизни человечества, с которым не может сравниться все им раньше пережитое. Недалеко время, когда человек получит в свои руки атомную энергию, такой источник силы, который даст ему возможность строить свою жизнь, как он захочет. Это может случиться в ближайшие годы, может случиться через столетие. Но ясно, что это должно быть.

Сумеет ли человек воспользоваться этой силой, направить ее на добро, а не на самоуничтожение? Дорос ли он до умения использовать эту силу, которую неизбежно должна дать ему наука?

Ученые не должны закрывать глаза на возможные последствия их научной работы, научного прогресса. Они должны себя чувствовать ответственными за последствия их открытий. Они должны связать свою работу с лучшей организацией всего человечества...»

Глава о геохимии радиоактивных элементов в его «Очерках геохимии» — одна из лучших.

Начатое 50 лет назад В. И. Вернадским изучение геохимической истории редких химических элементов способствовало созданию в нашей стране необходимых предпосылок к возможно более широкому использованию этих элементов в нашем социалистическом строительстве.

Мы неминуемо идем к полному вовлечению в промышленность всех химических элементов и даже искусственно получаемых. Поэтому современное значение геохимических исследований по распределению химических элементов в земной коре еще более возрастает. Оно обязывает нас искать законы этого распределения. У нас имеются к тому все возможности.





СОДЕРЖАНИЕ

МАТЕМАТИКА

ФИЗИКА, ГЕОФИЗИКА И АСТРОНОМИЯ

ХИМИЯ

СОДЕРЖАНИЕ



Печатается по постановлению Редакционно-издательского совета Академии Наук СССР

РИСО АН СССР № 2941. А-09383. Типографский заказ № 2990. Подписано к печати 18/IX 1947 г. Печ. л. $90^1/_2$ + 20 вклеек. Уч.-издат. л. $43^1/_4$. Тираж 3 300. Формат бумаги 82 × $108^1/_{16}$.

2-я типография Издательства Академии Наук СССР. Москва, Шубинский пер., д. 10

63

